Engineering inducible cell lines for recombinant Adeno-Associated Virus production

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE
UNIVERSITY OF MINNESOTA
BY

Zion Lee

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Advisor: Professor Wei-Shou Hu

August 2021

Acknowledgements

I am extremely grateful for the guidance and mentorship of my advisor, Professor Wei-Shou Hu. He has always made graduate student development his top priority, and never misses an opportunity to help us grow. I am constantly inspired by his commitment to collaborations, pioneering in the unknown, and prescient decision making. I have had a PhD experience that I could not have even dreamed of. Professor Hu has shaped me into the best scientist that I can be.

I want to express appreciation for all the great collaborators who have made our joint projects stimulating and rewarding. I thank Professor Daniel Schmidt, who helped us start our AAV work and guided our projects towards success and relevance in the field. I thank Professor Nikunj Somia for offering molecular biology insight into viral and mammalian gene expression engineering. I thank Professor Alon Herschhorn for our exciting work on developing stable cell lines for HIV and SARS-CoV-2 vaccine antigen production. And I would like to thank Professor Samira Azarin, Professor Ben Hackel, and Professor Nikunj Somia for taking the time to serve on my committee.

Next, I would like to mention the members of the Hu group who have made my time at the University of Minnesota truly special. I really appreciate all that Dr. Arpan Bandyopadhyay was to me: teacher, mentor, friend. Huge thanks also to Thu Phan, who has endured the struggles of graduate school with me since day one. A special acknowledgment to Dr. Chris Stach, who taught me so much about molecular biology and vector design. I also want to thank all the members of the Hu group, past and present, for the support and great times we were able to share: Dr. Yonsil Park, Dr. Dong Seong Cho, Dr. David Chau, Dr. Tung Le, Dr. Conor O'Brien, Dr. Sofie O'Brien, Dr. Jen One, Dr.

Kevin Ortiz-Rivera, Dr. Hansol Kim, Meghan McCann, Min Lu, Qian Ye, Yen-An Lu, and Janani Narayan.

A special thanks to the undergraduates who have helped make our projects possible. I thank Eesha Irfanullah for contributing her brilliant mind to achieve our grand goals. I thank Morgan Soukup for contributing her relentless work ethic to the work at hand. I thank Marina Raabe and Noelle Schumacker for showing interest and initiative toward independent research. It has truly been an honor to be a supervisor and mentor to you all and one of the most rewarding aspects of my graduate school experience.

Lastly, I want to thank my family. I thank my parents and my sister for full and absolute support without adding any pressure. I want to thank my son for being my happy cheerleader. And I want to thank my wife for sustaining me, encouraging me, and making daily sacrifices to get me through to the end.

Abstract

The recombinant adeno-associated virus (rAAV) gene therapy field has experienced landmark regulatory approvals in recent years by demonstrating high efficacy in treating monogenic diseases. There is therefore an immediate and increasing need for highly productive manufacturing platforms to generate large quantities of rAAV vectors. Current rAAV synthesis methods require multiple-component plasmid transfections or viral infections, which have increasing costs, variability, and technical challenges at large scales. To address the growing need for robust methods to produce and characterize rAAV vectors, we used synthetic biology tools to gain control over viral gene expression dynamics. First, we engineered a rAAV infectious titer assay cell which contains the AAV and helper genes necessary to replicate transduced genomes. Once exposed to small molecule inducers and a rAAV vector, the assay cell line performed similarly to the standard adenovirus (Ad) coinfection method in determining the infectious titer of the vector preparation. Omitting the use of Ad greatly simplifies the potency assay and reduces the variability of input reagents. Building on that success, we further engineered cells that not only have replicative capacity, but also packaging capacity and a latent copy of a rAAV genome. Small molecule induction allowed expression of the necessary viral genes to synthesize infectious rAAV vectors. The variation of inducer ratios enabled manipulation of vector quality, in terms of the fraction of DNA-containing particles. These synthetic cellular technologies address the needs of a reliable potency bioassay and a scalable and robust production platform for rAAV manufacturing.

Table of Contents

List of Tables

List of Figures

1 Introduction

Among vectors for human gene therapy, recombinant adeno-associated virus (rAAV) vectors are advantageous for their long-term efficacy, safety profile and serotype specificity. These therapies are able to dramatically change the lives of patients with previously untreatable diseases. The FDA approvals of Luxturna for retinal dystrophy in 2017 and Zolgensma for spinal muscular atrophy in 2019 have generated much excitement about the future potential of this modality (He, Urip, Zhang, Ngan, & Feng, 2021). With the accelerating pace of new clinical trials, it is becoming increasingly important to have reliable bioassays to assess product functionality and innovations to make production more scalable and robust. With the new modality comes unique manufacturing challenges distinct from those faced with recombinant protein therapeutics such as monoclonal antibodies. Many of these difficulties can be overcome through the infusion of synthetic biology in vector manufacturing technology.

Synthetic biology is a broad, expanding field consisting of many aspects. It involves the creation of new functionalities by learning from biological systems and engineering them towards a desired objective. Many recent advances have been in the areas of biomolecule sensing, synthetic genetic circuits, and biological elements responsive to external inputs (Kitada, DiAndreth, Teague, & Weiss, 2018). Inducible promoters activate gene transcription in response to inducers such as antibiotics, hormones, or even light (Kallunki, Barisic, Jäättelä, & Liu, 2019). Another mode of control is at the protein level, with many tools available for ligand-responsive protein degradation (T. Wu et al., 2020). These control tools, in conjunction with recent advances in transposon technology to easily

integrate large, multi-gene cassettes into mammalian cells (Balasubramanian, Wurm, & Hacker, 2016), create many opportunities to reinvent complex cellular systems.

Viral replication and packaging are complex processes in mammalian cells. They involve a multitude of viral and host factors which enhance or mitigate one another. The ultimate goal of a wild-type virus is to make more progeny virions, often completely overtaking the host cell and causing cell death. Viral vector manufacturing takes this complex process and perturbs it away from its optimum by separating vector components from viral genes. It is not surprising then that vector titer, quality and potency are diminished compared to the wild-type virus (Zeltner, Kohlbrenner, Clément, Weber, & Linden, 2010). Synthetic biology can be the tool used to push the system back closer to its optimal functioning point. Further, the tendency of viruses to kill their host cells requires their genes to be controlled if host cells stably harboring viral genes are desired.

1.1 Thesis organization

This thesis represents the application of synthetic biology to innovate processes in rAAV manufacturing. Chapter 2 details the construction of a replication-competent cell line that can amplify transduced genomes and thereby enable sensitive readout of a vector's potency. Chapter 3 details the construction of an exclusively inducible producer cell line which can generate infectious rAAV and enables manipulation of full particle content. Chapter 4 summarizes the work and provides directions for further development of the technology. In addition, the addendum, also called chapter 5, provides additional information into innovations in cell line development for discovering the most favorable epigenetic landing spots for transgene integration.

2 Development of an Inducible, Replication-Competent Assay Cell Line for Titration of Infectious Recombinant Adeno-Associated Virus Vectors

Reproduced from: Lee, Z., Lu, M., Irfanullah, E., Soukup, M., Schmidt, D., Hu, W. Development of an Inducible, Replication-Competent Assay Cell Line for Titration of Infectious Recombinant Adeno-Associated Virus Vectors. (In preparation)

2.1 Summary

An important quality attribute of a recombinant Adeno-Associated Virus (rAAV) as a therapeutic vector is its infectious titer. Current infectious titer assays mostly rely on coinfection of rAAV with a helper virus such as adenovirus, which requires helper virus preparation and introduces additional variability. Here we describe a stable infectious titer assay cell line into which was integrated the coding sequences for AAV Rep68 and adenovirus E4orf6 and DNA Binding Protein (DBP) under the control of inducible promoters. The Rep68 protein expression was further modulated by a ligand-responsive destabilization domain. In parallel assays the cell line gave comparable titer to that obtained using a classical adenovirus coinfection method. The cell line was also used to titer vectors of multiple AAV serotypes. This cell line has the potential to serve as an effective and robust tool for product quality evaluation.

2.2 Introduction

Infectious titer is a critical indicator of the potency of a gene delivery product and is one of the most important quality attributes of a rAAV vector. But rAAV therapies are

typically dosed based on physical titer, in vector genomes per mL (vg/mL). Another measure often used to quantify rAAV is the total particle titer, measured in capsids per mL. Since there is not a universal relationship among physical titer, infectious titer, and total particle titer, it is crucial to characterize the ratios between them. Achieving a particle-to-infectious unit ratio close to unity, which has been observed with wild-type AAV, represents the optimal scenario for a gene therapy vector (Zeltner et al., 2010). Yet for rAAV it was reported that the ratio between genome containing particles and infectious particles, called infectivity, was highly variable between production methods (Gao et al., 1998). In addition, different production platforms yield varying ratios of genomes to capsids, also called the full-to-empty particle ratio (Merten, 2016). The vector genome copy can be measured by quantitative PCR (Werling, Satkunanathan, Thorpe, & Zhao, 2015), while the total particle concentration can be measured by ELISA of capsid proteins (D. Grimm et al., 1999). Both are easily carried out in a typical laboratory.

The infectious titer, in contrast, is commonly assessed using coinfection with a helper virus such as adenovirus (Ad) to provide helper functions for transduction and/or replication. For example, the replicative center assay (RCA) infects stable *rep*/*cap*-expressing HeLa cells with rAAV and Ad, then transfers cells onto a nylon membrane and finally detects replicated rAAV genomes by fluorescent or radiolabeled probes (Salvetti et al., 1998; Zolotukhin et al., 1999). In another method, *rep*/*cap*-expressing HeLa cells seeded in a 96-well plate are infected with replicated dilutions of rAAV along with adenovirus, and the wells are assayed for vector genome replication by qPCR (Lock et al., 2010; Zhen, Espinoza, Bleu, Sommer, & Wright, 2004). The theoretical dilution at which half of the replicates would be replication-positive is used to calculate the 50% tissue

culture infective dose ($TCID_{50}$). Infectious titer can also be measured by observing the number of cells expressing transgene after coinfection of rAAV with adenovirus (Xiao, Li, & Samulski, 1998). We will use the common notation that genome-detecting assays are reported in infectious units (IU) and transgene-detecting assays are reported in transducing units (TU).

Ad coinfection requires additional Ad stock production and characterization. Coinfection methods may also introduce the potential bias of adenovirus-mediated endosomal disruption (François et al., 2018), which would result in over-estimation of rAAV infectious titer. As a complex bioreagent, Ad is a contributor to assay variability. The infectious titer assays already showed high variability in performance, especially between different characterization sites (Lock et al., 2010). A more robust infectious titer assay, preferably without using helper virus coinfection, will be welcome.

AAV genome replication utilizes host cellular DNA replication machinery, but also relies on a subset of its own genes and some Ad genes. The large Rep78 or Rep68 protein processes AAV ITRs to unwind and create necessary nicks which are used to initiate the replication of AAV genomes (Gonçalves, 2005; Stracker et al., 2004). The adenovirus DNA binding protein (DBP) binds to single-stranded genomes and enhances replication processivity (Ward, Dean, O'Donnell, & Berns, 1998), making it a requirement for efficient AAV replication (Matsushita et al., 1998). The E4orf6 protein of Ad mediates second-strand synthesis of AAV genomes (Ferrari, Samulski, Shenk, & Samulski, 1996), a critical step in converting AAV genomes to double-stranded forms for transcription and DNA replication.

Previous attempts at generating stable cell lines expressing AAV and Ad genes have been hindered by the cytotoxicity of the AAV *rep* gene products (Qiao, Wang, Zhu, Li, & Xiao, 2002; Schmidt, Afione, & Kotin, 2000) and the Ad E2A and E4 gene products (Qiao, Li, Skold, Zhang, & Xiao, 2002). Our approach uses synthetic, inducible promoters to control the expression of AAV replication and Ad helper genes. We show that controlled expression of these genes supports efficient AAV genome replication for high infectious titer assay sensitivity.

2.3 Methods

2.3.1 Vector construct

The replication module consists of *rep* and helper genes controlled by inducible promoters and the necessary transactivators. Under control of the GeneSwitch™ promoter (Invitrogen) were the adenovirus E4orf6 and DNA binding protein (DBP) sequences linked by the P2A self-cleaving peptide. The GeneSwitch™ transactivator (GSTA) transcriptional unit was included from the pSwitch vector (Invitrogen). Under control of the TetON promoter (Clontech) was the DD-mCherry-Rep68 fusion protein sequence, DD representing the mutant FKBP12 destabilization domain (Takara Bio Inc.). The final transcriptional unit included the reverse tetracycline transactivator (rtTA3) linked to the hygromycin-resistance protein sequence by the T2A self-cleaving peptide. All four transcriptional units were assembled into a piggyBac transposon vector (ATUM).

2.3.2 Cell culture and transfection

HEK293 cells (Cell Biolabs, Inc.) were cultured in Dulbecco's Modified Eagle Medium (Gibco) supplemented with 10% Fetal Bovine Serum (Gibco) and antibiotic-antimycotic (Gibco). For assay cell construction, cells were transfected with an equal

amount of replication module plasmid and Leap-In transposase mRNA (ATUM) using Lipofectamine 3000 (Invitrogen). Transfected cells were cultured in media containing 200 µg/mL Hygromycin B (Millipore Sigma) for two weeks to obtain a stable cell pool. The pool was seeded into 96-well plates at a dilution of 0.5 cells per well to obtain single cell clones. Each well was visually inspected under a microscope to ensure cells were derived from a single cell. Clones were expanded and cryopreserved.

Triple transfection of pAAV-CAG-GFP (Addgene #37825), pHelper (Cell Biolabs, Inc.) and pAAV-RC2 (Cell Biolabs, Inc.) was conducted on HEK293 cells to generate rAAV2-GFP. Cells were lysed by three freeze-thaw cycles (-80°C/37°C) and treated with benzonase to result in a crude preparation. Purified vectors of AAV serotypes 2, 6, and 8 were generated using the same pAAV-CAG-GFP plasmid by the University of Minnesota Viral Vector and Cloning Core. Viral vectors were purified with sucrose cushion ultracentrifugation resulting in titers ranging from 1.13×10^{12} to 8.69×10^{13} vg/mL.

2.3.3 Flow cytometry assays

RM4 cells were seeded at 1.5×10^5 cells per 24-well in basal media with 5 µg/mL doxycycline and 10 nM mifepristone. Varying dilutions of rAAV2-GFP were added to the wells the following day. After a two-day incubation period, cells were detached and analyzed on a BD LSR II flow cytometer for GFP fluorescence. For immunofluorescence detection of GFP expression, detached cells were fixed with 4% paraformaldehyde, permeabilized with 0.2% Triton X-100, incubated with mouse anti-GFP (Sigma G6539), then incubated with goat anti-mouse F(ab')2 fragment, Alexa Fluor 647 conjugate (Invitrogen A-21237). Then cells were analyzed on a BD LSR II flow cytometer for Alexa Fluor 647 fluorescence.

2.3.4 Infectious titer plate assay

The plate assay for infectious titer used quantitative PCR-based detection of replication events. Cells were seeded at 2.6×10^4 cells per 96-well and induced with 5 µg/mL doxycycline and 10 nM mifepristone. Varying dilutions of rAAV2-GFP stock suspension were added to the wells in replicates of ten the following day. After a two-day incubation period, cellular DNA was harvested by QuickExtract™ solution (Lucigen) and quantitative PCR was run using primers against GFP. The mean cycle number and standard deviation of ten uninfected wells were used to define the negative range, and wells with cycle numbers three standard deviations below that mean were considered positive (Zhen et al., 2004). The Spearman-Kärber method was used to determine the 50% tissue culture infective dose, $TCID_{50}$ (Kärber, 1931; Ramakrishnan, 2016).

2.4 Results

2.4.1 Infectious titer assay development

The assay cell line, RM4, was constructed by integrating the replication module construct into HEK293 cells (Figure 2-1A) and screening among several single cell clones (Figure 2-2). Induction by mifepristone activates the GeneSwitch™ promoter, resulting in transcription of the Ad E4orf6 and DBP coding sequences. Induction by doxycycline activates the TetON promoter, resulting in transcription of the DD-mCherry-Rep68 fusion protein. DD is a mutant FKBP12 destabilization domain which upon tagging to a protein marks the protein for degradation unless stabilized by the Shield1 ligand (Banaszynski, Chen, Maynard-Smith, Ooi, & Wandless, 2006). Tagging Rep68 with DD and placing it under TetOn inducible promoter control minimized its possible cytotoxic effects. Transduction of rAAV2-GFP into uninduced RM4 cells resulted in only a few cells

expressing GFP, and the fluorescence intensity was low (Figure 2-1B). In contrast, RM4 cells induced to express Rep68 and helper proteins were able to support the replication of exogenously transduced rAAV genomes and express the cargo GFP proteins. This was evident by strong GFP expression in many RM4 cells transduced with rAAV2-GFP particles (Figure 2-1C).

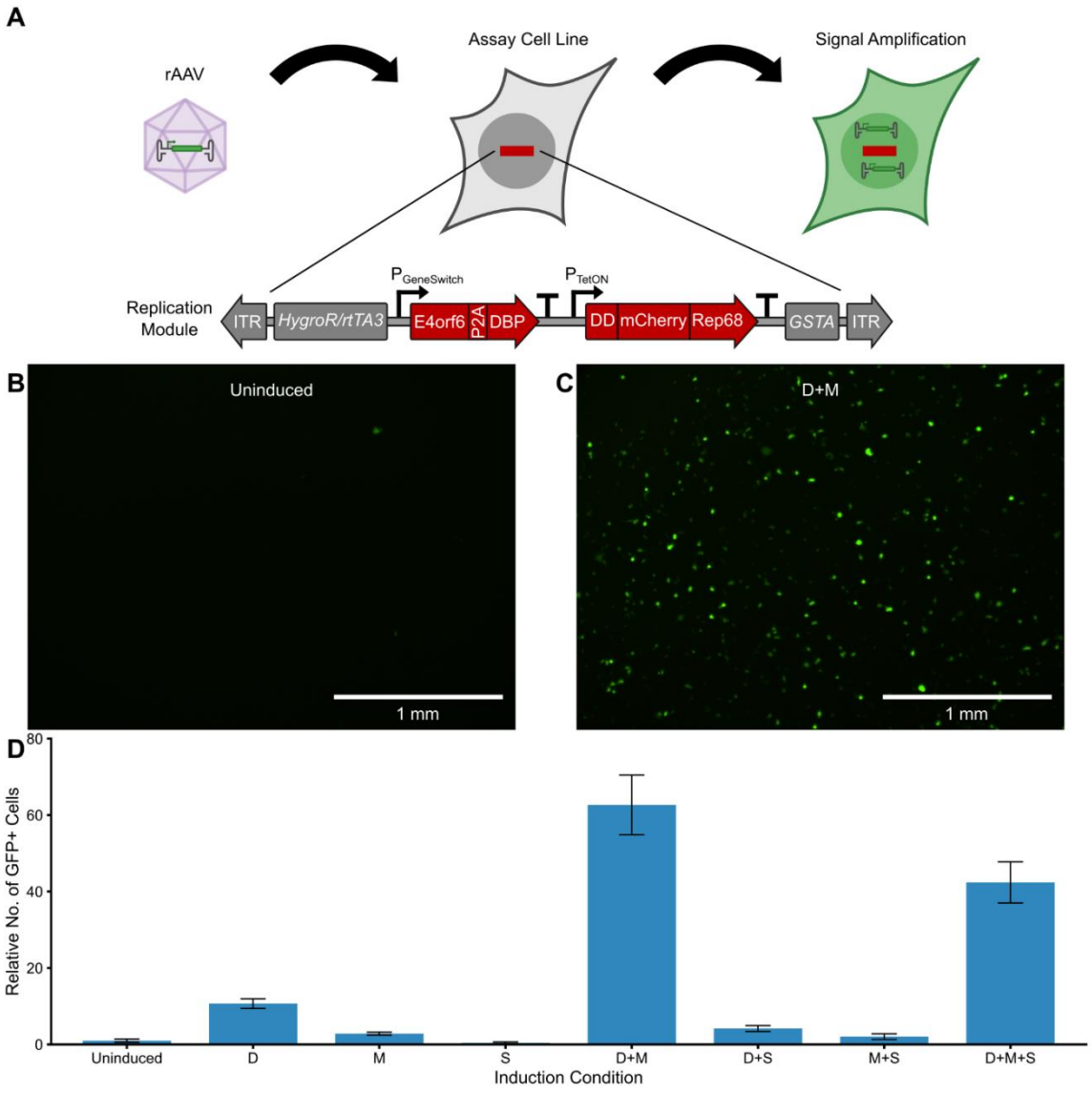

Figure 2-1: Design and construction of an inducible rAAV assay cell line.(A) Replication-competent cell line conceptual schematic showing high vector protein expression upon transduction. Inset shows integrated construct design with its inducible elements. (B) GFP

fluorescence image of RM4 cells transduced with rAAV2-GFP without induction. (C) GFP fluorescence image of RM4 cells transduced with rAAV2-GFP with doxycycline and mifepristone induction. (D) Relative infectious titer of RM4 cells upon transduction with rAAV2-GFP and induced with permutations of three molecules of interest. All values relative to apparent titer using uninduced cells. Data represents mean and standard deviation of four image-based cell counts. D=doxycycline, M=mifepristone, S=shield1.

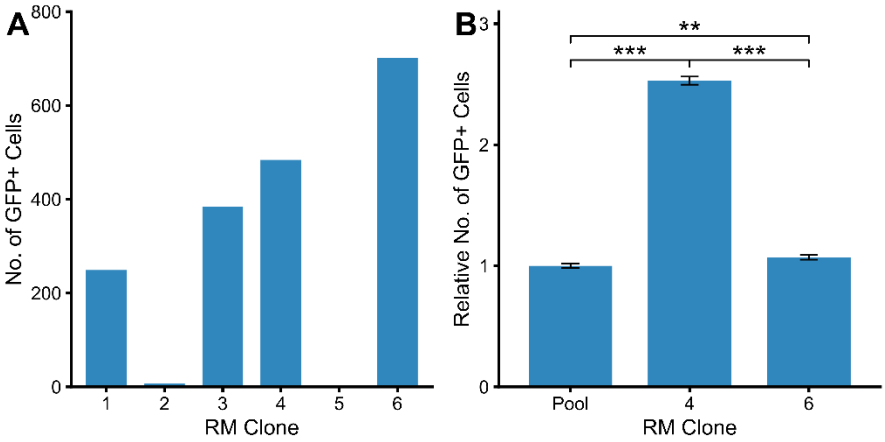

Figure 2-2: RM clone screening. (A) Round 1 screening of six R2 clones by adding a fixed amount of virus to induced cells. Transduction events counted by ImageJ. (B) Round 2 screening, comparing the two best clones with the pool under controlled seeding density. Data represents mean and standard deviation of three independent experiments. Significance assessed by Welch's *t*-test. **$p<0.01$, ***$p<10^{-5}$. Clone 4 performed much better than the pool.

The induction conditions strongly affected the number of transduced cells. Without any inducers, few cells were GFP positive. Inducing with both doxycycline and mifepristone gave the highest number of transduced cells, while moderate results were seen with doxycycline or mifepristone alone (Figure 2-1D). In all subsequent experiments RM4 cells were induced with doxycycline and mifepristone. In all combinations of inducers, stabilizing Rep68 with Shield1 reduced the apparent titer compared to the case without Shield1. This is consistent with an earlier finding that reduction of large *rep* protein expression enhanced AAV replication (Li, Samulski, & Xiao, 1997).

2.4.2 Assay cell line benchmarking using flow cytometry

Infectious titration of rAAV using RM4 cells was benchmarked against the traditional method of Ad coinfection. Flow cytometry measurement showed that the two assays gave similar fractions of GFP-positive cells (Figure 2-3A). A small fraction of cells in the RM4 assay had a very high, near saturation level of GFP expression, which was not seen in the Ad coinfection assay, likely due to viral genome amplification occurring in RM4 cells. The apparent titer of the rAAV2-GFP viral prep calculated from the fraction of GFP-positive cells was very similar between the RM4 assay ($5.59 \pm 0.15 \times 10^8$ TU/mL) and Ad coinfection assay ($6.66 \pm 0.16 \times 10^8$ TU/mL) (Figure 2-3B).

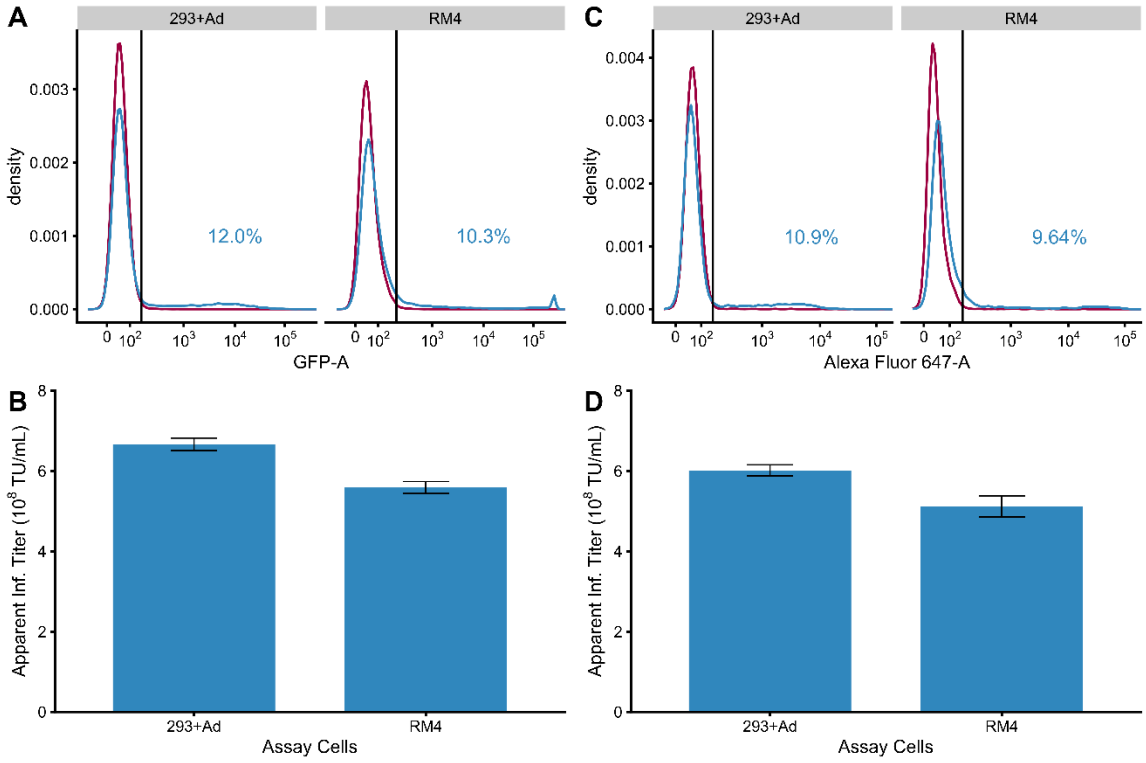

Figure 2-3: Benchmarking apparent infectious titer of RM4 cells with a traditional method. (A) Flow cytometry population histograms of GFP-intensity for HEK293 cells coinfected with adenovirus and induced RM4 cells. The blue curve represents cells transduced with rAAV2-GFP, and the red curve represents untransduced cells. (B) Apparent infectious titer by adenovirus infection or RM4 clone, measured by GFP. Data represents mean and standard deviation of three replicates. (C) Flow cytometry population histograms of Alexa

Fluor 647-intensity for immunostained HEK293 cells coinfected with adenovirus or induced RM4 cells. The blue curve represents cells transduced with rAAV2-GFP, and the red curve represents untransduced cells. (D) Apparent infectious titer by adenovirus infection or RM4 clone, measured by Alexa Fluor 647-tagged Anti-GFP antibody. Data represents mean and standard deviation of three replicates.

In clinical applications, the payload gene will not be a reporter but a therapeutic protein. Immunostaining of the cargo protein was used to detect the transduction of rAAV. GFP protein in rAAV2-GFP-infected RM4 cells was detected by immunostaining with a mouse anti-GFP antibody and an Alexa Fluor 647-tagged anti-mouse secondary antibody and benchmarked against the Ad coinfection assay with the same immunostaining procedure. Flow cytometry measurement of immunostained RM4 cells and Ad coinfected cells gave similar fractions of Alexa Fluor 647-positive cells (Figure 2-3C). As with the GFP-based method, RM4 cells showed a subpopulation with very high magnitude of Alexa Fluor 647 fluorescence. Again, the RM4 assay gave an apparent infectious titer of the rAAV2-GFP preparation ($5.12 \pm 0.26 \times 10^8$ TU/mL) very similar to one obtained by the Ad coinfection assay ($6.01 \pm 0.14 \times 10^8$ TU/mL) (Figure 2-3D).

2.4.3 Plate assay for infectious titration

We next compared fluorescence-based assays using RM4 with a qPCR-based $TCID_{50}$ assay. For the $TCID_{50}$ assay, rows of ten wells of induced RM4 cells were transduced with dilutions of rAAV2-GFP such that at a limiting dilution, wells will probabilistically receive one (or more) infectious units, or none at all. Replicated genomes were detected using qPCR, and the positive wells detected were shown as a plate-view map (Figure 2-4A). The percent positive wells from each row were plotted against the dilution factor, and the Spearman-Kärber method was used to calculate the infectious titer as 6.0 ±

1.8×10^8 $TCID_{50}$/mL across three replicates (Figure 2-4B). The $TCID_{50}$ assay titer was thus similar to the GFP fluorescence and immunofluorescence assay titers.

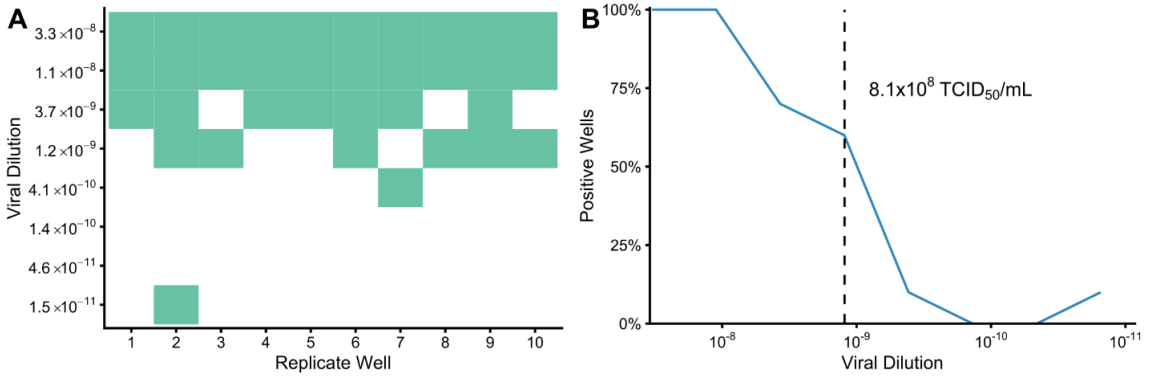

Figure 2-4: $TCID_{50}$ method of assaying infectious titer using RM4 cell line. (A) Visual plate view of a representative limiting-dilution $TCID_{50}$ infectious titer assay. Green squares indicate wells which were positive for vector genomes via quantitative PCR. Experiment was performed on RM4 cells in 96-wells. (B) Calculation of apparent titer according to the Karber method.

2.4.4 rAAV Infectivity

Two independent preparations of the same GFP-bearing genome were available to compare infectivity ratios. Vector genome titers were measured using qPCR and the ratio to infectious titer measured by GFP fluorescence was calculated to be the infectivity. One preparation, which represented a relatively crude extraction of rAAV2-GFP vector, had a VG:TU ratio of 175 (Table 2-1). Another preparation was purified by ultracentrifugation and had a VG:TU ratio of 86,447 (Table 2-1).

Table 2-1: Infectivity of two rAAV2 preparations. The VG:TU infectivity ratio of two rAAV2 preparations are shown here. The first preparation was a crude preparation of rAAV2-GFP generated in house, used for Figure 1. The second preparation was a highly purified vector used for Figures 2 and 3.

rAAV2-GFP preparation	Infectivity (VG:TU)
1	175 ± 25
2	86,447 ± 29,193

2.4.5 Titration of different AAV serotypes

The rAAV virus stocks used in the assays described above were of serotype 2. HEK293 cells, from which RM4 were derived, express surface receptors for other AAV serotypes including serotypes 6 and 8 which are commonly used in gene therapy studies. RM4 cells were used to determine the infectious titer of rAAV-GFP of serotypes 6 and 8 (Table 2-2). As with serotype 2, assessment of serotypes 6 and 8 were comparable between RM4 assays and the Ad coinfection assays.

Table 2-2: Infectious titers of additional serotypes. Infectious titers of rAAV6 and rAAV8 harboring a GFP transgene, measured using flow cytometry analysis of HEK293 cells coinfected with Ad and RM4 assay cells. Data represents mean ± standard deviation of three replicates.

	Apparent Infectious Titer (TU/mL)	
AAV Serotype	**293+Ad**	**RM4**
AAV6	$9.24 \pm 0.21 \times 10^6$	$1.06 \pm 0.05 \times 10^7$
AAV8	$1.26 \pm 0.17 \times 10^7$	$1.66 \pm 0.21 \times 10^7$

2.5 Discussion

Infectivity is a critical determinant of the functionality of an rAAV vector for gene therapy. A robust quantification method of infectivity is important to manufacturing process optimization and characterization. We constructed an rAAV assay cell line, RM4, to enable a rapid and sensitive measurement of infectious titer. These cells expressed AAV replication and Ad helper elements upon induction with small molecules to allow them to replicate the vector genome upon infection of rAAV. The amplification of infected rAAV genomes allowed for sensitive detection of cargo protein expression or presence of vector genomes and easy enumeration of rAAV-infected cells for titer quantification. The

infectious titers determined using RM4 was comparable to those determined using the conventional adenovirus coinfection method for serotypes 2, 6, and 8. The RM4 based titer assay eliminates helper virus from the protocol making the method simpler, safer and less susceptible to variability in helper virus preparation.

Using flow cytometry, we quantified the fraction of transduced RM4 cells or Ad coinfected HEK293 cells as indicated by GFP fluorescence or immunostaining of GFP. Our flow cytometry measurements had good intra-assay repeatability between replicates, indicated by relative standard deviations below 6%. RM4 cells reported infectious titers 16% and 15% lower than the classical Ad coinfection method for GFP fluorescence and immunofluorescence assays, respectively. Although the differences appeared significant, bioassays typically have high degrees of intermediate and inter-assay variability, corresponding to assays conducted in separate runs and separate facilities, respectively (J. W. Lee et al., 2006). Bioassays for potency especially have high levels of overall variability, with standard deviation reported up to 101% of the mean in rAAV2 transgene expression assays for infectious titer (Lock et al., 2010). Therefore, the slightly lower infectious titer with RM4 cells compared to Ad coinfected cells is relatively minor compared to range of the expected bioassay variability. Nonetheless, minor differences may have been influenced by the previously mentioned endosomal disruption mediated by Ad which can lead to over-estimation of rAAV titers (François et al., 2018). Another possible contributing factor of the difference could have been a cellular host response. rAAV replication is known to trigger DNA damage response proteins such as ATM (Schwartz, Carson, Schuberth, & Weitzman, 2009), which in turn inhibit rAAV transduction (Zentilin, Marcello, & Giacca, 2001). Despite these known phenomena potentially driving

differences in titer, the RM4 cells obtained a final titer relatively similar to the current standard method of Ad coinfection, demonstrating their utility.

Immunostained titers were measured to be slightly lower than the respective GFP titers using the corresponding assay cells. These slight differences could have been the result of non-specific binding of primary or secondary antibodies in the untransduced controls (see Figure 2-5), which may have increased the gating threshold for calling positive events. However, these differences were less than 10%, which is similar to sources of variability such as cell counting uncertainty (Cadena-Herrera et al., 2015).

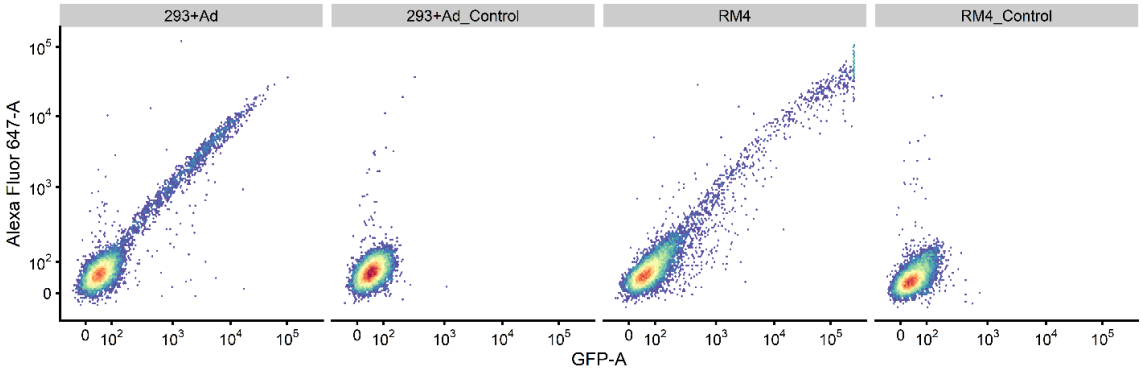

Figure 2-5: Two-dimensional cytograms of immunostained assay cells. Flow cytometry results showing Alexa Fluor 647 (anti-GFP) signal and GFP signal of adenovirus-coinfected HEK293 cells (left pair) and doxycycline- and mifepristone-induced RM4 cells (right pair). Cells transduced with rAAV2-GFP (left of pair) are shown along with their untransduced controls (right of pair). Most Alexa Fluor 647- and GFP-positive cells were also positive for the other, lying completely on the diagonal region of the plot. This demonstrates any cargo gene encoding a protein of interest can be assayed using these cells with a simple immunostaining assay.

RM4 cells were also used in a $TCID_{50}$ quantitative PCR-based assay. Using the same rAAV sample, remarkable consistency was observed between flow cytometry-based and qPCR-based assays despite the methods being completely orthogonal. The final ratio of $TCID_{50}$-derived and GFP fluorescence-based titers yielded a $TCID_{50}$:TU ratio of 1.07. The $TCID_{50}$:TU ratio close to unity contrasts with previous studies which performed both

types of assays and found a $TCID_{50}$:TU ratio of 8.6 (Lock et al., 2010) and 125 (François et al., 2018). The cited ratios may have been higher because of their use of adenovirus coinfection, longer assay time (72 hours), and host cell origin (HeLa). Yet we believe the good agreement between flow cytometry-based titers and $TCID_{50}$ titer is a strength, demonstrating the consistency and reliability of a stable, inducible cell line.

Using the RM4 assay cell line, the infectivity of two different AAV preparations were found to be in between 10^2 and 10^5 VG:TU (Table 2-1), which is in the range reported previously using HEK293 assay cells (Gao et al., 1998; Mohiuddin et al., 2005). It was encouraging that the RM4 assay cells enabled assessment of vector potencies over nearly three orders of magnitude. One potential reason for the difference in rAAV potency between preparations could be different degrees of aggregation, which have been observed in some rAAV vectors following DNase treatment (Dobnik et al., 2019). In addition, post-translational modifications to AAV capsids and vector genome methylation were observed to vary based on rAAV production platform (Rumachik et al., 2020). Finally, the method of purification has also been shown to impact infectivity (El Andari & Grimm, 2021).

The RM4 assay cells were also shown to be effective at measuring infectious titers of rAAV serotypes 2, 6 and 8. It was reported that HEK293 cells are not efficiently transduced by serotypes 4, 5, and 9 (Dirk Grimm et al., 2008). A previous attempt to increase pan-serotype infectivity by overexpressing the “universal” AAVR receptor resulted in improvement in only some serotypes (McDougald et al., 2019). It may be of interest to increase the sensitivity of the assay cell line toward inefficient serotypes by expressing specific receptors or receptor processing genes. Those cells would be an important step towards a universal infectious titer assay cell line.

2.6 Conclusion

Until now, infectious titer assays have required adenovirus coinfection, representing tedious and advanced laboratory techniques. These RM4 assay cells allow for rapid and sensitive titration of viral particles using induction with small molecules. Doxycycline and mifepristone are inexpensive inducers which come with much lower source variability than a complex biological agent such as adenovirus. The cell line represents a product of the growing demand for high quality assays to assess vector product functionality.

3 Construction of a rAAV Producer Cell Line through Synthetic Biology

Reproduced from: Lee, Z., Lu, M., Irfanullah, E., Soukup, M., Hu, W. Construction of a rAAV Producer Cell Line through Synthetic Biology. (In preparation)

Z. Lee performed producer cell construction and characterization. M. Lu performed inducer response experiments for producer cells and packaging cell construction and characterization.

3.1 Summary

The complexity of recombinant adeno-associated viruses (rAAV) poses a challenge to the creation of stable cell lines as a scalable platform for rAAV production. Here we describe a template for a transfection-free, helper virus-free rAAV producer cell line using synthetic biology. Three modules were integrated into HEK293 cells including an rAAV genome and multiple inducible promoters controlling all necessary AAV Rep, capsid, and helper coding sequences. The synthetic cell line yielded roughly 500 vector genomes per cell and was stable for at least 30 generations. Independent control over replication and packaging genes allowed for manipulation of the full capsid content, reaching up to 73%. Robust rAAV genome amplification was observed, indicating potential areas of improvement in capsid protein expression. The achievement represents an approach that enables highly controllable production of rAAV at large scale.

3.2 Introduction

Recombinant AAV production is much more complex than traditional biologics manufacturing. First, the product is a nanoparticle consisting of roughly sixty capsid proteins at a 1:1:10 ratio of VP1:VP2:VP3 proteins (Gonçalves, 2005) encapsidating a single-stranded DNA molecule of up to 5 kilobases in size (Z. Wu, Yang, & Colosi, 2010). Second, viral replication involves the interaction of multiple viral and host factors which accelerate cell death. The large *rep* proteins Rep78 or Rep68 nick AAV genomes in the Internal Terminal Repeats (ITRs) (Tullis Gregory & Shenk, 2000), facilitating replication by the host cell DNA polymerase. Adenoviral single-stranded DNA-Binding Protein (DBP) improves processivity of DNA replication (Ward et al., 1998) and is also involved in forming sub-nuclear viral replication centers (Hidalgo et al., 2016; Weitzman, Fisher, & Wilson, 1996). Adenoviral E4orf6 protein has many functions, including export of viral mRNA from the nucleus, blockage of p53 activity, and stimulation of AAV second-strand synthesis (Leppard, 1997). Adenoviral VA RNA are non-coding RNAs which bind to and inactivate PKR, subduing the cellular anti-viral response of protein synthesis arrest (Vachon & Conn, 2016). Capsid assembly is catalyzed by the Assembly Activating Protein (AAP) and localized to the nucleolus (Sonntag, Schmidt, & Kleinschmidt, 2010). Finally, the small *rep* proteins Rep52 or Rep40 are required for insertion of single-stranded AAV genomes into pre-formed capsids (King, Dubielzig, Grimm, & Kleinschmidt, 2001). More viral factors are continuously being discovered, such as the membrane-associated assembly protein (MAAP) which add further complexity to the viral life cycle (Ogden, Kelsic, Sinai, & Church, 2019). The concert of biological reactions complicate efforts to make manufacturing platforms efficient yet consistent.

The most common platform for rAAV vector generation is the transfection of plasmids encoding the therapeutic cassette flanked by AAV ITRs, the AAV *rep* and *cap* genes, and adenovirus helper genes *E2A*, *E4*, and *VA RNA* (Clément & Grieger, 2016). This method generates vectors in a short timeline, but has limitations at large scales, including GMP plasmid preparation, transfection efficiency, and reagent cytotoxicity (Srivastava, Mallela, Deorkar, & Brophy, 2021). Alternative technologies typically integrate some of the necessary components into the cellular genome or helpervirus genome. Stable integration of AAV *rep/cap* and rAAV genome into HeLa (Chadeuf et al., 2000; Clark, Voulgaropoulou, Fraley, & Johnson, 1995) and A549 (Farson et al., 2004) cells resulted in clones which produced rAAV upon infection with adenovirus. Herpes simplex virus (HSV) was identified as an alternative helpervirus, inspiring a production system where suspension Baby Hamster Kidney cells are co-infected with a recombinant HSV containing the rAAV genome and another recombinant HSV containing AAV *rep/cap* genes (Thomas et al., 2009). In the OneBac system, stable AAV *rep* and *cap* sequences in Sf9 insect cells are induced upon infection with a recombinant baculovirus which harbors the rAAV genome (Mietzsch et al., 2013). However, these methods possess other challenges, such as helpervirus removal in purification (Clément, Knop, & Byrne, 2009), GMP helpervirus preparation (Wright, 2008) and helpervirus genetic stability (Kohlbrenner et al., 2005).

For decades the biologics industry has achieved great success by creating stable cell lines which can be cultured in suspension at high densities. The main barrier to achieve stable cell lines for AAV production has been the toxicity of *rep* (Schmidt, Afione, & Kotin Robert, 2000) and helper genes (Qiao, Li, et al., 2002). The key to navigating around this challenge has been synthetic biology, which has seen much recent development especially

in mammalian applications. Many previous efforts have shown effective control of the most toxic protein, Rep78, using the dexamethasone-inducible mouse mammary tumor virus promoter (Hölscher et al., 1994), the heavy metal-inducible mouse metallothionein promoter (Yang, Chen, & Trempe, 1994), and the doxycycline-inducible Tet-On promoter (Chen, Clark, Chen, Schnepp, & Johnson, 2002). Similar promoters have been used to control Adenovirus helper proteins DBP and E4orf6 (Brough, Lizonova, Hsu, Kulesa, & Kovesdi, 1996; Kovesdi & Hedley, 2010). To control the levels of toxic Rep52 genes, a "dual splicing switch" was designed, representing a synthetic intron with a Cre-excisable array of polyadenylation signals (Qiao, Wang, et al., 2002). An early effort to integrate all necessary components for AAV production resulted in cell lines which were not stable (Qiao, Li, et al., 2002). Recently, a fully inducible *rep*, *cap*, and Helper system has been developed, demonstrating the necessity of synthetic biology in stable AAV production (Coronel et al., 2021).

Being able to produce rAAV is not the only objective for a stable cell line. The resulting platform should generate high quality vectors, especially in the aspect of full capsid content. Patient immune responses pose a great challenge for rAAV therapies (Manno et al., 2006; Mingozzi & High, 2013) and having excess empty capsids will only increase the antigenic load without delivering the therapeutic payload. It has traditionally been the burden of downstream purification efforts to separate the full capsids from empty capsids. Ultracentrifugation methods using CsCl or iodixanol gradients are effective in separation based on density differences but are not easy to use at large scale (El Andari & Grimm, 2021). Ion exchange chromatography is a more scalable method that can efficiently separate full particles from empty particles based on isoelectric point (Tomono

et al., 2018). Nevertheless, low full capsid content and variability between batches put additional burden on downstream processes.

In this study, we constructed stable producer and packaging cell lines of rAAV using synthetic biology tools. We were able to demonstrate control over viral genes and subsequently rAAV titer and full capsid content. This novel strategy enables scalable manufacturing of rAAV with control over vector quality in the production phase.

3.3 Methods

3.3.1 Vector Design

The rAAV2-GFP genome from pAAV-CAG-GFP (Addgene #37825) was cloned into a piggyBac transposon backbone (ATUM, Inc.) with a puromycin resistance gene driven by a phosphoglycerate kinase promoter, yielding the rAAV genome module. AAV Rep68 sequence was synthesized by Integrated DNA Technologies and the Rep40 start codon ATG was mutated to GGG by PCR, as previously reported (Chejanovsky & Carter, 1989). Rep68 was tagged with mCherry and a mutant FKBP12 destabilization domain (Takara Bio) (Banaszynski et al., 2006). The resulting fusion protein was cloned downstream of the TetON promoter (Takara Bio). The adenovirus E4orf6 and DBP sequences were PCR amplified from the pHelper plasmid (Cell Biolabs, Inc.) and joined by a P2A linker. The resulting CDS was cloned downstream of the GeneSwitch promoter (ThermoFisher Scientific). The Rep68 and helper transcriptional units, as well as the GeneSwitch transactivator gene, were cloned into a piggyBac transposon backbone with a hygromycin resistance gene linked via T2A to the reverse Tet transactivator, yielding the replication module. AAV *cap* CDS was PCR amplified from pAAV-RC2 (Cell Biolabs, Inc.) and modified to omit the native intron and mutate the VP1 start codon to ACG, as

described previously, and termed VP123 (Urabe, Ding, & Kotin, 2002). The VP123 sequence was linked to AAV Rep52 via the ECMV IRES and placed under the control of the CumateSwitch promoter (System Biosciences). The transcriptional unit was cloned into a piggyBac transposon backbone with a blasticidin resistance gene linked via T2A to the Cym Repressor, yielding the packaging module.

3.3.2 Cell Culture

HEK293 cells (Cell Biolabs, Inc.) were cultured in Dulbecco's Modified Eagle Medium (DMEM, Gibco) supplemented with 10% Fetal Bovine Serum (Gibco) and 1X antibiotic-antimycotic (Gibco). Stable cell lines were generated by co-transfection of transposon vectors with Leap-In transposase mRNA (ATUM, Inc.) at a 1:1 DNA:RNA ratio. After three days of outgrowth, cells were passaged into media containing 2 µg/mL puromycin (Invivogen), 200 µg/mL hygromycin B (MilliporeSigma), and/or 10 µg/mL blasticidin (Invivogen). After two weeks of selection, single-cell cloning was performed in 96-wells by limiting dilution. RM4 infectious titer assay cells were previously developed (submitted manuscript) by integration of the replication module into HEK293 cells. Producer cell lines Pf3 and Pf6 were generated by integration of the genome, replication, and packaging modules into HEK293 cells. Packaging cell line RP6 was generated by integration of the packaging module into RM4 assay cells.

Infectious titer assays were conducted by seeding RM4 cells at 1.5×10^5 cells per 24-well with 5 µg/mL doxycycline (Sigma Aldrich) and 10nM mifepristone (ThermoFisher Scientific). rAAV preparations were added to the cultures one day after seeding, and infectious titers were measured two days post transduction. For rAAV production via producer cells, cells were seeded at 3.75×10^5 cells per 6-well with 5 µg/mL doxycycline,

10nM mifepristone, and 30 µg/mL cumate (Sigma Aldrich), unless noted otherwise. Cells were harvested four days post seeding. For rAAV production via packaging cells, cells were seeded at 7.5×10^5 cells per 6-well with 5 µg/mL doxycycline, 10nM mifepristone, and 30 µg/mL cumate. One day post seeding, cells were transduced with seed rAAV2-GFP virus, and cells were harvested four days post seeding. Intracellular DNA and RNA were harvested using the Quick-DNA/RNA Miniprep (Zymo Research D7001) and quantified by qPCR and qRT-PCR using standard methods.

3.3.3 rAAV Preparation and Characterization

Harvested cells were lysed by three cycles of freeze/thaw at -80°C/37°C and treated with ≥250 units/mL Benzonase (Millipore Sigma E1014) for one hour. Benzonase activity was quenched by adding 2mM EDTA, and cell debris was removed by centrifugation. For physical titration, proteinase K (Macherey-Nagel, 740506) was added to release DNA, and ssDNA was isolated by column purification (Zymo Research). Quantitative PCR was performed on purified genomes with pAAV-CAG-GFP as a reference standard using primers against the GFP coding sequence. Capsid titer was measured using an A20 antibody ELISA kit (Progen PRATV).

3.4 Results and Discussion

To construct a rAAV producer cell line, three genetic modules were integrated into the genome of HEK293 cells: the rAAV genome, replication module, and packaging module (Figure 3-1A). The replication module consisted of adenovirus E4orf6 and DBP linked by P2A under the GeneSwitch promoter control and AAV Rep68 tagged by DD and mCherry under the TetON promoter control. The DD was a destabilization domain that could be stabilized upon the addition of the Shield1 ligand. Because Rep68 was suspected

to be the most cytotoxic viral factor, the design used both the DD and the TetON promoter, which showed the lowest degree of leaky expression (Figure 3-2). In the packaging module, intron-less AAV *cap*, called VP123, and AAV Rep52, were controlled by the CumateSwitch promoter. Since *cap* transcripts were known to be expressed at the highest levels (Mouw Matthew & Pintel David, 2000; Stutika et al., 2016), they were placed under the CumateSwitch promoter, which conferred the highest level of induced expression (Figure 3-2). Because it was reported that the sequence downstream of an IRES experiences attenuated translation (Mizuguchi, Xu, Ishii-Watabe, Uchida, & Hayakawa, 2000), Rep52 was linked to VP123 via an IRES sequence since small Rep transcripts are not as highly abundant as *cap* transcripts (Mouw Matthew & Pintel David, 2000). The three vectors were stably integrated via piggyBac transposase into HEK293 cells using three separate selection markers.

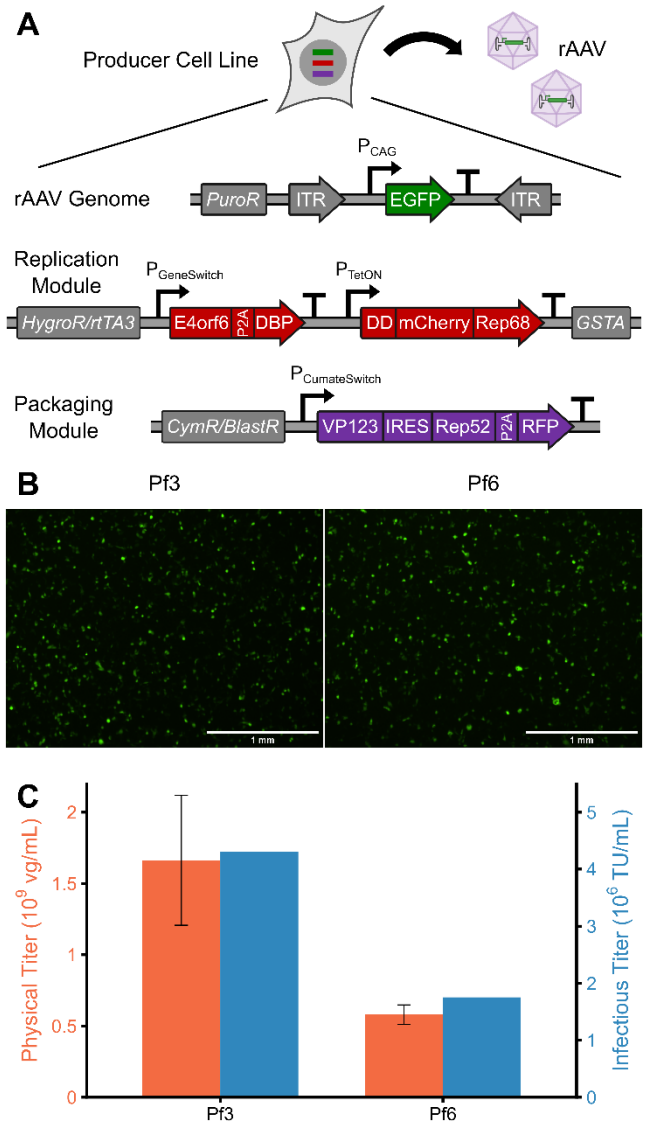

Figure 3-1: Vector design and producer clone characterization. (A) A synthetic rAAV producer cell line and its three AAV genetic elements. GSTA denotes the GeneSwitch transactivator gene; rtTA3 denotes the reverse Tet transactivator gene; CymR denotes the Cym Repressor; DD denotes the Destabilization Domain. (B) GFP fluorescence images of infectious titer assay cells transduced with rAAV2-GFP produced by Pf3 or Pf6 clones. (C) Virus titer obtained through the induction of Pf3 and Pf6 cells. Physical titers, measured by qPCR, showing the mean and standard deviation of three qPCR wells. Infectious titers measured by flow cytometry analysis of rAAV-transduced assay cells.

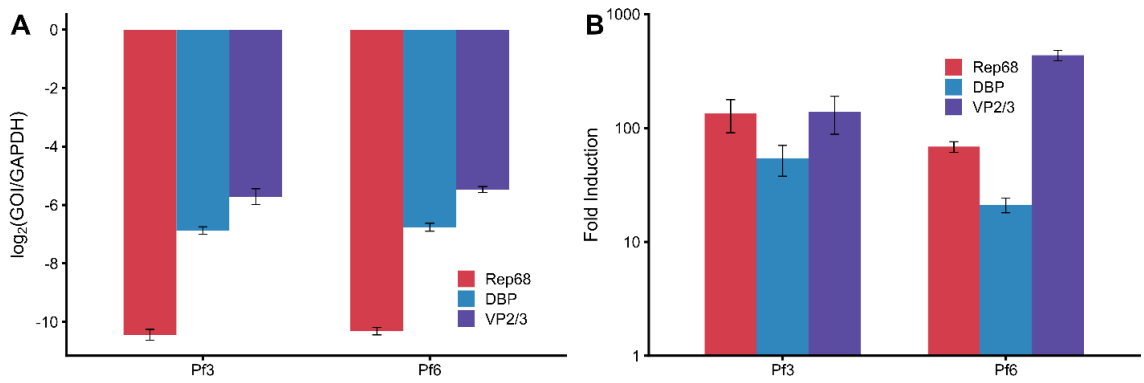

Figure 3-2: Inducibility of viral transcripts in Pf3 and Pf6 producer cells. (A) Uninduced Pf3 and Pf6 transcript expression levels of viral genes and (B) fold-induction.

Two producer cell clones, Pf3 and Pf6, were isolated. Upon induction, both clones were able to produce infectious rAAV vectors (Figure 3-1B), as evidenced by GFP fluorescence in RM4 infectious titer assay cells developed previously (manuscript in preparation). The physical titer, in vector genomes per mL, and the infectious titer, in transducing units per mL were measured for clones Pf3 and Pf6 (Figure 3-1C). Pf3 had higher titers, hence Pf3 was used in further experiments. Clonal differences may have arisen from variance in transcriptional activity of synthetic genes, which is influenced by the epigenomic contexts of transgene integration sites (Z. Lee, Raabe, & Hu, 2021), and transgene copy numbers (Joe Carver et al., 2020). In addition, clones could have had different levels of innate immune response factors which trigger inhibitory reactions to viral processes (Muhuri et al., 2021).

The productivity of Pf3 cells was responsive to inducer concentrations. The infectious rAAV yield increased with increasing doxycycline concentration (Figure 3-3A) and increasing cumate concentration (Figure 3-3B). These results demonstrated a strong control over viral genes in the uninduced states, and robust induction of viral genes in the induced states (see also Figure 3-2). The ratio of the doxycycline and cumate

concentrations had an impact on the particle titer and the full capsid content (Figure 3-3C). At low doxycycline and high cumate levels, particle titers were high but full capsid contents were low. Conversely, with high doxycycline and low cumate levels, particle titers were low but full capsid content was high, approximately 73%. This demonstrated the ability to modulate full capsid content using independent control of viral genes. Using the triple transfection method, we obtained approximately 35% full capsid content. In a previous study, full capsid contents of multiple runs were below 20% (D. Grimm et al., 1999), while in another study, full capsid content was approximately 50% (Zeltner et al., 2010). It is clear that the full capsid content is highly variable and unpredictable, but by adjusting ratios of chemical inducers, Pf3 cells were driven to increase this important vector quality attribute.

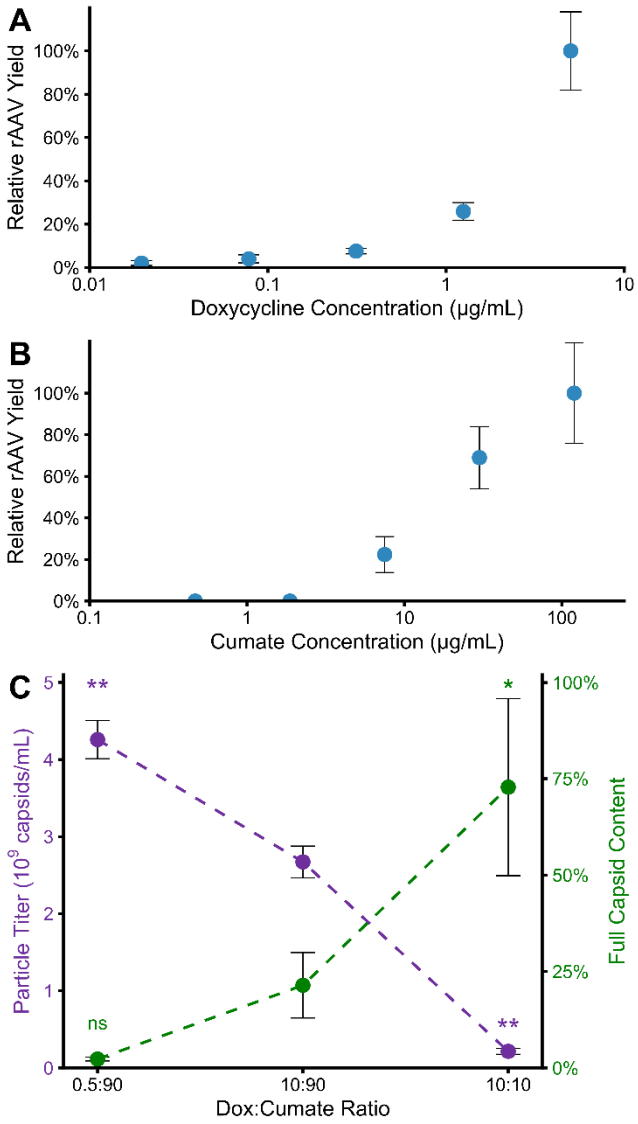

Figure 3-3: Response of Pf3 cells to inducer concentrations. (A) Effect of doxycycline concentration on infectious rAAV yield, measured by infectious titer assay cells. The relative titer was in reference to the titer obtained with the highest doxycycline concentration. Mifepristone and cumate concentrations were kept at 10nM and 30µg/mL, respectively. Data represents mean and standard deviation of cell counts from twelve images. (B) Effect of cumate concentration on infectious rAAV yield, measured by infectious titer assay cells. The relative titer was in reference to the titer obtained with the highest cumate concentration. Data represents mean and standard deviation of cell counts from twelve images. (C) Effect of varying dox:cumate ratios on particle titer and full capsid content. Presented numbers for Dox and Cumate represent the actual concentration of the inducers used, in µg/mL. Data represents mean and standard deviation of three independent replicates. Significance assessed by Welch's *t*-test in comparison to the 10:90 Dox:Cumate ratio condition. *p<0.05, **p<0.01, ns = not significant.

New technologies should be benchmarked against the current standard methods. To that aim, the specific physical and particle titers of Pf3 producer cells were compared to those values of the widely used triple-transfection method (Figure 3-4A). Both values in Pf3 cells (506 vg/cell and 703 capsids/cell) were roughly two orders of magnitude lower than the triple-transfection method. While specific titers may be lower, theoretically volumetric titers can be similar or higher due to recent advances in perfusion technology which allows cell densities up to 10^8 cells/mL (Schwarz et al., 2020). Still, there may be opportunities to further improve Pf3 or equivalent cells in terms of vector design and process optimization. When comparing gene expression between the two methods, Pf3 cells had roughly 4-fold lower large Rep and DBP/E2A transcript levels but similar VP2/3 levels, VP2/3 being the most important *cap* transcript isoform (Figure 3-4B). The VP2/3 transcript levels being similar contrasted with the lower particle titer observed in Figure 3-4A. A possibility is that protein levels in Pf3 did not reach the levels of triple transfection, or potentially there were enough proteins but a lack of assembly into capsid particles.

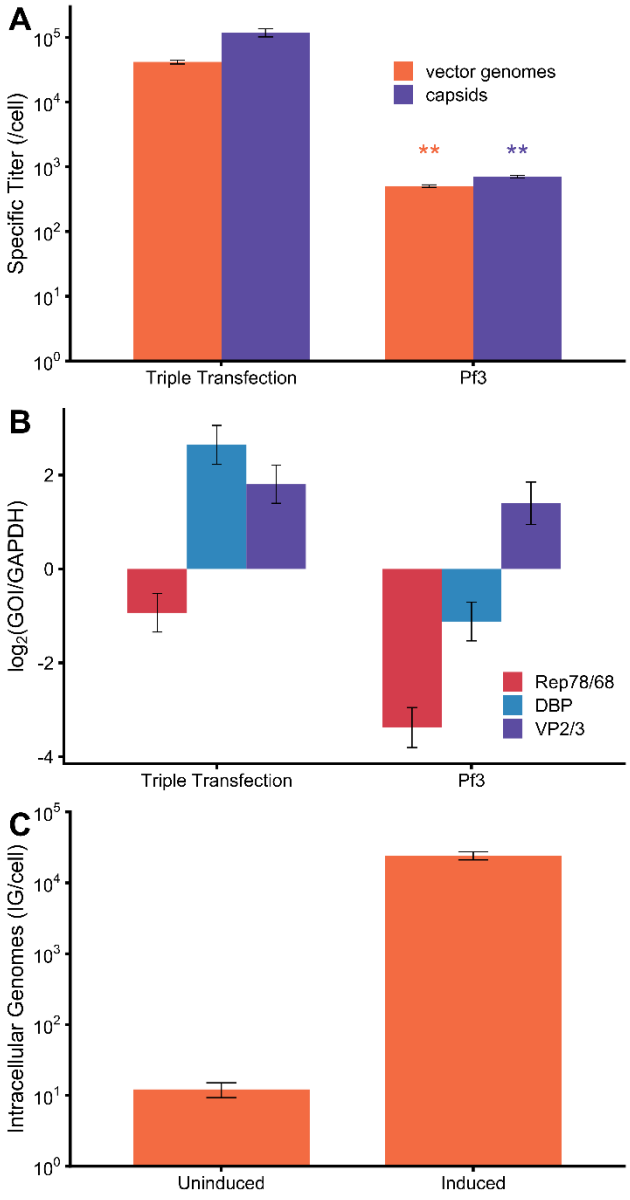

Figure 3-4: Comparison of Pf3 performance with triple transfection rAAV production method. (A) Specific physical (vector genomes per cell) and particle (capsids per cell) titers of triple transfection and Pf3 production methods. Data represents mean and standard deviation of replicate qPCR wells and capsid ELISA wells. Significance assessed by Welch's *t*-test in comparison to the Triple Transfection condition. **p<0.01. (B) Transcript expression of viral genes at the end of production. Data represents mean and standard deviation of triplicate qPCR wells. (C) Total intracellular genomes (packaged and unpackaged) in uninduced and induced Pf3 cells. Data represents mean and standard deviation of triplicate qPCR wells.

One notable feature omitted from the synthetic cell lines that is usually present in the triple transfection production method was adenovirus VA RNA. Adenoviral VA RNAs have been reported to be expressed at extremely high levels (10^8 copies/cell) (Vachon & Conn, 2016) and they have been shown to induce the type I interferon innate immune response (Yamaguchi et al., 2010). Potential cytotoxicity was one reason why we chose to omit integration of VA RNA. Additionally, the target of VA RNA, PKR, is typically triggered by double-stranded RNA, which we thought would not be present at high levels in our synthetic cell lines due to the replacement of viral promoters with inducible promoters. The discrepancy between the capsid mRNA levels and capsid titers in Pf3 cells, may indicate a potential avenue for improvement in integrating the Ad VA RNA. Encouragingly, doxycycline-inducible expression of a VA RNA was previously shown to be feasible in a stable cell line (Machitani et al., 2011).

To assess if replication could also be an explanation to the disparity in physical titers observed, the total intracellular genomes of Pf3 cells were measured in uninduced and induced states (Figure 3-4C). Uninduced copy numbers were low, representing the basal copy number of stably integrated rAAV genomes. Upon induction, genome copy number reached an excess of 10^4 genomes per cell, which was in the same order of magnitude as the triple transfection physical titers. Thus, it appears that the level of genome replication in Pf3 cells may not be the limiting factor despite the lower transcript levels of large Rep and helper genes seen in Figure 3-4B. This was consistent with the finding that reduced large Rep protein expression using an ACG inefficient start codon did not reduce titers but rather led to increased titers (Li et al., 1997). Pf3 producer cells and RP6 packaging cells also showed reduced viral titers when Shield1 was added to stabilize Rep68,

suggesting that Rep68 protein levels were sufficient (data not shown). Thus, even though the capability to increase Rep68 concentrations was implemented in our design, it was kept in the destabilized state for improved rAAV production in all experiments presented. Yet as more knowledge is gained about the temporal requirements of large Rep proteins, the stabilizable control of Rep68 remains a tool which can be utilized to fine-tune protein levels for optimal rAAV production.

Due to the cytotoxicity of Rep and Helper genes, assessment of the stability of the cell line was regarded as crucial. In commercial manufacturing, cells must be sufficiently stable in order to faithfully perform their functions after expansion from cryovials to large scale, high density bioreactors. Pf3 cells were passaged in selection media for 30 doublings and the physical titers were compared to the early generation of cells (Figure 3-5). The physical titer remained in the same range, and on a per-cell basis did not decrease over this time. This was encouraging to see that performance could be maintained through long-term culture. A previous study which aimed to integrate all AAV and helper components did not obtain stable cell lines presumably because of uncontrolled expression from E2A and E4 (Qiao, Li, et al., 2002). A cell line constitutively expressing E2A has been documented (Zhou, O'Neal, Morral, & Beaudet, 1996), so the cytotoxic gene was likely E4. Taken together, it appears that at minimum, all presented AAV factors – large Rep, small Rep, and capsid proteins and Ad factor E4orf6 should have gene expression control in some form to obtain stable AAV producer cell lines.

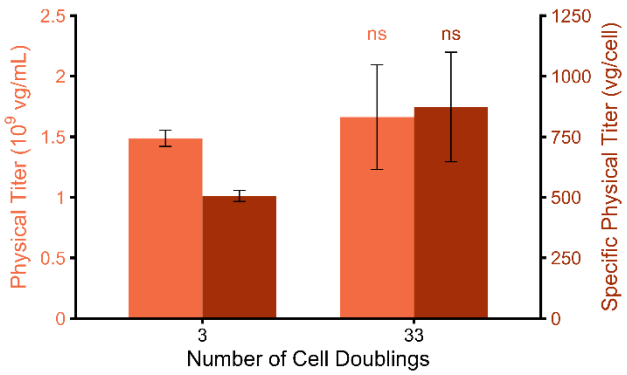

Figure 3-5: Pf3 cell line stability. Physical titer and specific physical titer measured for Pf3 cells induced at early and late generation numbers. Data represents mean and standard deviation of replicate qPCR wells. Significance assessed by Welch's *t*-test in comparison to the 3 Doublings condition. ns = not significant.

A variation of the rAAV producer cell line concept was that of a packaging cell line, which did not possess an integrated copy of the rAAV genome, but still contained the replication and packaging modules. In these cells, a rAAV genome could be supplied by seed virus infection or plasmid transfection (Figure 3-6A). A packaging cell clone, RP6, was induced and infected with seed rAAV2-GFP at 175 vg/cell to produce more virus. The resulting vectors were transduced into assay cells and successful transduction events were seen (Figure 3-6B). In contrast, very few GFP-positive assay cells were observed when HEK293 cells were used as a mock control. This confirmed that the vectors measured were a result of new synthesis, not the original seed virus carrying through to the final preparation. The infectious titer of RP6 and HEK293 cells were calculated, showing a clear difference in the amount of virus produced in each case (Figure 3-6C). The overall infectious titer of RP6 cells was lower than the Pf3 producer cells of Figure 3-1. Possible causes may have been insufficient dose of seed virus and/or integrated vector copy numbers.

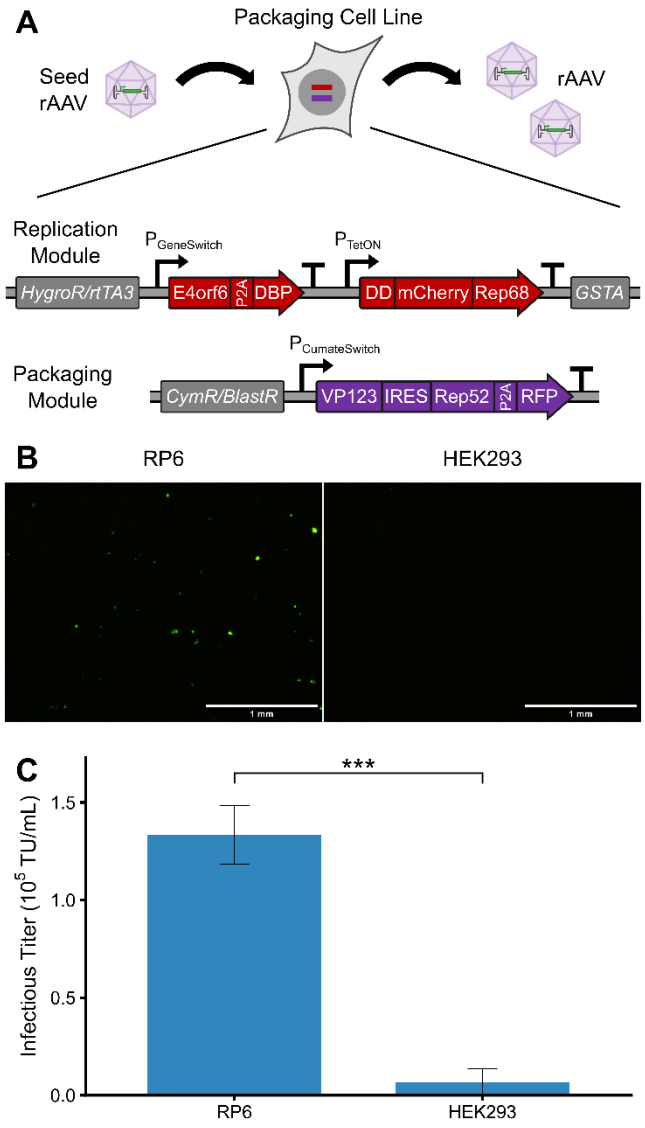

Figure 3-6: Vector design and packaging clone characterization. (A) A synthetic rAAV packaging cell line and its two AAV genetic elements. GSTA denotes the GeneSwitch Transactivator gene; rtTA3 denotes the Reverse Tet Transactivator gene; CymR denotes the Cym Repressor; DD denotes the Destabilization Domain. (B) GFP fluorescence images of infectious titer assay cells transduced with rAAV2-GFP from packaging cell clone RP6 or HEK293 cells. (C) Virus infectious titer of induced RP6 cells or HEK293 cells, measured by infectious titer assay cells. Data represents mean and standard deviation of cell counts from six images. Significance assessed by Welch's *t*-test. ***$p<10^{-5}$.

Stable cell lines for rAAV production have been elusive for many years. Here we demonstrated the extensive use of synthetic biology in mammalian cells to achieve fully inducible production of complex recombinant virus particles. Low basal transcription and

high induced transcription was achieved, and opportunities still exist in boosting protein translation for higher vector titers. Independent control over replication and packaging modules enabled unprecedented manipulation over the vector full particle content. Stable cell lines are the ultimate goal for scalable manufacturing, and the present work exemplifies the crucial role of synthetic biology to engineer next-generation gene therapy manufacturing platforms.

4 Summary and Concluding Remarks

Gene therapy using rAAV vectors is a therapeutic modality that is generating much excitement due to clinical successes. In order to maintain the rate of growth of the field, vector manufacturing technology must advance to address the challenges associated with making a complex bio-nanoparticle product and assuring its potency and consistency. It was seen how various production and purification methods led to drastically different vector infectivities. Current methods of measuring infectious titer had very high inter-assay variability, which makes troubleshooting and process understanding difficult. Vector design and production platform also yield vectors with different full capsid ratios, but very little can be done to improve those ratios other than extensive downstream purification operations. These potential pitfalls must be addressed for the continued success of rAAV therapies.

This thesis described a synthetic biology approach that led to two cell-based technologies which bear good news for the industry. The infectious titer assay cell line removed the need for helper virus coinfection, eliminating the bioreagent as a source of variability. And yet it performed comparably to the standard helper virus coinfection method, demonstrating its ability to promote second-strand synthesis and/or genome replication. The producer cell line integrated in full all the necessary components for rAAV synthesis. This makes feasible a cell line that can be grown to very large volumes and high densities. Infectious vectors were produced, and by varying the level of inducer molecules, the full capsid content was pushed to very high levels. The cells demonstrated the level of influence that can be exerted during rAAV production to adjust vector titer and quality.

Opportunities for further development of the technologies exist. Viral gene expression programs have specific dynamics that maximize their replicative potential. These can be studied by means of time-series transcriptomics and proteomics to determine the levels of viral factors at intervals post infection. The use of synthetic biology allows for the ability to mimic the transcript and protein expression profiles to achieve replication closer to a wild-type infection. But AAV genes are not the only available targets for optimization. AAV interacts with the host cell for control over cellular machinery. The ease of host cell genome engineering using technologies such as CRISPR/Cas9 allow knock out or activation of host cell factors which may restrict AAV replication. Finally, the cell culture environment is different from the natural environment that AAV would normally infect a host. That provides opportunities to study how the AAV replication and packaging processes are impacted by the cell culture conditions. Process parameters such as temperature, pH, and media formulation can have drastic impacts on cell physiology and metabolism. With the plentiful tools available for improvement, the future holds promise for rAAV vector production.

5 Addendum: Epigenomic features revealed by ATAC-seq impact transgene expression in CHO cells

Reproduced with permission from: Lee, Z., Raabe, M., and Hu, W. Epigenomic features revealed by ATAC-seq impact transgene expression in CHO cells. *Biotechnology and Bioengineering*. 2021; 118: 1851– 1861. https://doi.org/10.1002/bit.27701

5.1 Summary

Different regions of a mammalian genome have different accessibilities to transcriptional machinery. The integration site of a transgene affects how actively it is transcribed. Highly accessible genomic regions called super-enhancers were recently described as strong regulatory elements that shape cell identity. Super-enhancers were identified in Chinese Hamster Ovary (CHO) cells using Assay for Transposase-Accessible Chromatin Sequencing (ATAC-seq). Genes near super-enhancer regions had high transcript levels and were enriched for oncogenic signaling and proliferation functions, consistent with an immortalized phenotype. Inaccessible regions in the genome with low ATAC signal also had low transcriptional activity. Genes in inaccessible regions were enriched for remote tissue functions such as taste, smell, and neuronal activation. A lentiviral reporter integration assay showed integration into super-enhancer regions conferred higher reporter expression cells than insertion into inaccessible regions. Targeted integration of an IgG vector into the *Plec* super-enhancer region yielded clones that expressed the immunoglobulin light chain gene mostly in the top 20% of all transcripts with the majority in the top 5%. The results suggest the epigenomic landscape of CHO cells can guide the

selection of integration sites in the development of cell lines for therapeutic protein production.

5.2 Introduction

Since the dawn of recombinant biologics manufacturing, cell culture process titers have increased by several orders of magnitude. Much of that advance can be attributed to enhanced process engineering and high productivity of producing cell lines. To attain high productivity, the producing cell must express the transgene to a high transcript level and possess other cellular characteristics associated with high levels of protein secretion. These important traits must also be retained over a large number of cell divisions. High transcript levels of transgene have classically been achieved through co-amplification of a transgene and a selectable marker such as dihydrofolate reductase or glutamine synthetase (Kingston, Kaufman, Bebbington, & Rolfe, 2002). Yet the amplification process prolongs the timeline of cell line development and may introduce chromosomal aberrations (S. J. Kim & Lee, 1999; Wurm & Wurm, 2017), increasing the risk of cell-line instability. Site-specific recombination using loxP/Cre (Kito, Itami, Fukano, Yamana, & Shibui, 2002), FRT/FLP (Huang et al., 2007), and att/Bxb1 (Inniss et al., 2017) hold promise to shorten the timeline of cell line development. These methods integrate transgenes into pre-established "landing pads" which may confer high transgene expression. With advances in genome engineering tools such as CRISPR/Cas9 (Cong et al., 2013), landing pads can be rationally chosen based on bioinformatic analysis.

The integration location of a transgene affects its expression (Barnes, Bentley, & Dickson, 2003; Wallrath & Elgin, 1995). An ideal target integration site should have the local context that favors stable and highly active transcription. A class of potentially

favorable genomic neighborhoods is the super-enhancer. Super-enhancers are clusters of enhancers which were noted to have higher influence on transcription than regular enhancers (Sengupta & George, 2017; Thandapani, 2019). Many super-enhancers are associated with master regulator transcription factors that regulate the expression of genes in control of cell identity (Whyte et al., 2013). In cancer, increased expression of many oncogenes were attributed to super-enhancers which were not found in healthy tissues (Peng & Zhang, 2018; Sengupta & George, 2017). In some cases, the regulatory activity of super-enhancers can be co-opted to other genes via translocations in a process known as enhancer hijacking. Some examples include translocations of the MYC oncogene to both Ig-associated and non-Ig-associated super-enhancer loci (Walker et al., 2014), and the *GFI1B* oncogene to the *BARHL1*/*DDX31* and *PRRC2B* super-enhancer loci (Northcott et al., 2014). Super-enhancers may have similar effects on transgenes that integrate in their region of influence.

In contrast to highly active super-enhancers, other chromatin regions have repressive effects and may negatively affect the expression of integrated transgenes. The expression of the *white* transgene in *Drosophila* decreased when juxtaposed with heterochromatin (Wallrath & Elgin, 1995). Five distinct chromatin states were later described in *Drosophila*, with three states corresponding to traditional heterochromatin or repressive chromatin. These regions covered nearly three-quarters of the genome and had low transcriptional activity (Filion et al., 2010). These features pose a challenge to high transgene expression.

Given the potential influence of epigenomic features on the expression of integrated transgenes, we sought to better characterize super-enhancers and inaccessible regions in

CHO cells. ATAC-seq has been used to identify super-enhancers (Semenkovich et al., 2016). ATAC-seq offers a wide breadth of information by not focusing on a particular DNA binding protein, but rather capturing any event which displaces nucleosomes from DNA such as transcription factor binding and transcription initiation. ATAC involves relatively straight forward laboratory procedures and is easily accessible for many labs. Here, ATAC-seq was used to identify super-enhancers in CHO cells. Integrating transgenes near these loci and away from inaccessible regions has the potential to facilitate the construction of high producing cell lines of recombinant proteins.

5.3 Materials and Methods

5.3.1 Cell Culture

CHO-K1 cells (ATCC CCL-61) and derived cells were cultured in Ham's F-12K (Kaighn's) Medium (Gibco) with 10% Fetal Bovine Serum (Gibco) at 37°C and 5% CO_2. 2C10 cells, an IgG producer with a single copy of IgG heavy and light chain genes integrated in its genome were previously established using lentiviral transduction (O'Brien et al., 2018).

5.3.2 Next-gen Sequencing Analysis

The ATAC transposition reaction as previously established (Buenrostro, Wu, Chang, & Greenleaf, 2015) was carried out on fifty-thousand cells in exponential growth phase. Transposed DNA was purified using a MinElute PCR cleanup kit (Qiagen). After library preparation, >50 million paired end reads were obtained by an Illumina HiSeq 2500 instrument. Adapter sequences and low-quality base calls were trimmed using trimmomatic. Trimmed reads were mapped to the mitochondrial genome, then all

unmapped reads were mapped to the *Cricetulus griseus* genome (Vishwanathan et al., 2016) using Bowtie2 with parameters -N 1 -L 20 -i S,1,0.50 -D 20 -R 3 -I 38 -X 2000.

Peaks were called using MACS2 with parameters: -p 0.00001 --nomodel --shift -100 --extsize 200 --bdg. Peaks -1kb to +100bp of transcription start sites were excluded using HOMER. Neighboring peaks separated by less than 12.5kb were stitched together using Bedtools. Read coverage in stitched peaks was computed and background coverage from CHO-K1 whole genome sequencing (Xu et al., 2011) in the same region were subtracted to give the ATAC signal. Peaks were ranked by ATAC signal, the inflection point of the ATAC signal vs. rank curve was set as the cutoff, and the peaks above the cutoff were taken as super-enhancers (Whyte et al., 2013). Inaccessible regions were identified by using Bedtools complement which returned all genomic regions not covered by peaks, then filtering for those at least 100kb in size.

RNA sequencing data was analyzed as previously described (O'Brien et al., 2018).

5.3.3 Targeted Integration using CRISPR/Cas9

Synthetic messenger RNA for eSpCas9(1.1) was *in vitro* transcribed using the mMESSAGE mMACHINE T7 Ultra kit (Invitrogen) and sgRNA was *in vitro* transcribed using the MEGAshortscript T7 Transcription kit (Invitrogen). RNA was purified using the MEGAclear Transcription Clean-Up kit (Invitrogen). Donor plasmids were constructed by Type IIs assembly of BFP/tk backbone vector, IgG-dGFP expression cassette, and two 500bp homology arm gBlocks (Integrated DNA Technologies) using the SapI enzyme. CHO-K1 cells were electroporated with Cas9 mRNA, sgRNA, and donor plasmid using a Bio-Rad Gene Pulser XCell Electroporation system with 4mm cuvettes and the settings 300V, 15ms, square wave.

Cells were passaged for two weeks to dilute out excess donor plasmid, and subsequently cells were sorted on a BD FACSAria II Cell Sorter with a 100μm nozzle. GFP-positive, BFP-negative cells were sorted into 96-well plates at one cell per well. After four days, the cells were treated with 80μM ganciclovir for HSVtk-mediated killing. Cells were expanded and assayed for expression of IgG mRNA using the Cells-to-CT 1-Step Power SYBR Green kit (Invitrogen).

5.4 Results

5.4.1 Super-enhancers identified by ATAC-seq

ATAC-seq was performed on CHO-K1 and 2C10, a CHO-K1-derived single-copy IgG producer (O'Brien et al., 2018). ATAC-seq peaks were stitched and rank-listed by background-subtracted read coverage to give ATAC signal. The stitched peaks separated into two regions: a small subset with rapidly rising ATAC signal and a large subset of relatively low signal (Figure 5-1A). Those with signal higher than the inflection point in the curve were considered to be putative super-enhancers, (Whyte et al., 2013) with 731 and 1124 being identified for CHO-K1 and 2C10, respectively. The mean read coverage of the 50kb regions centered on super-enhancer regions was higher over a broad region than non-super-enhancer regions (Figure 5-1B).

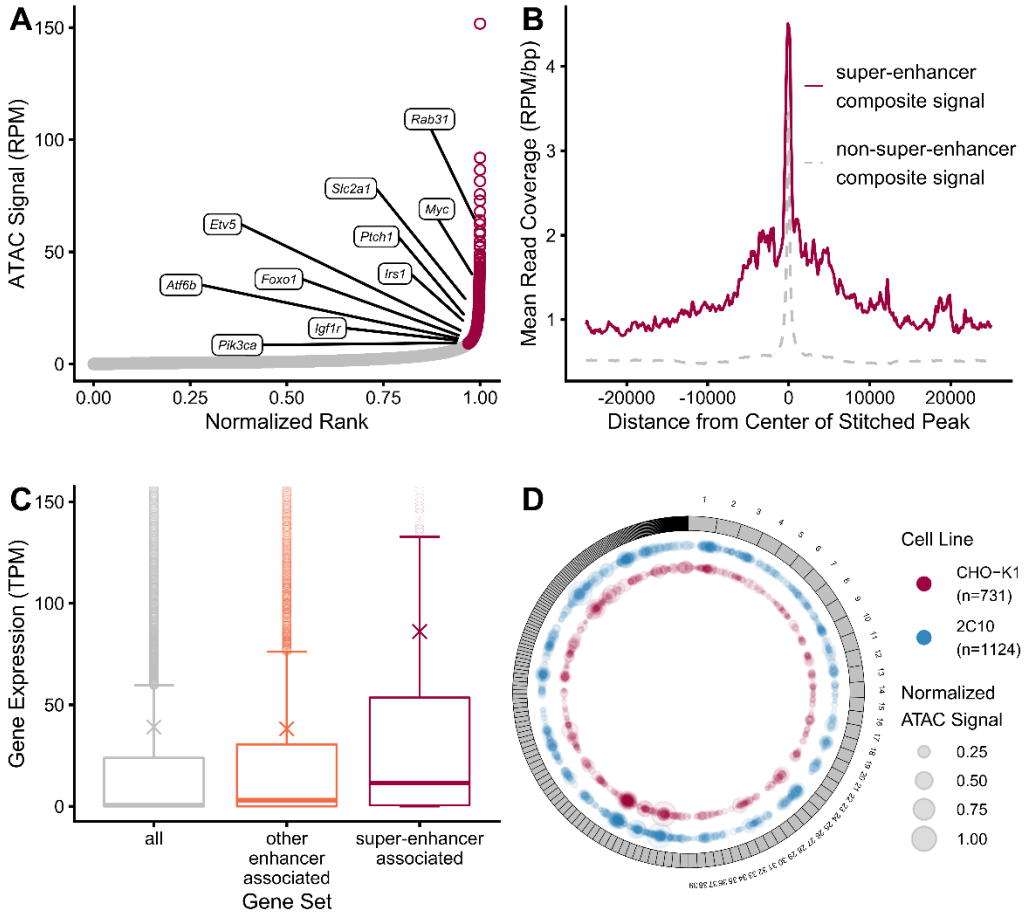

Figure 5-1: Super-enhancers in CHO and their associated genes. (A) Stitched ATAC-seq peaks in CHO-K1 cells ranked by background-subtracted read coverage, given in reads per million (RPM). Points above the infection point indicate stitched peaks designated as super-enhancers. Selected genes associated with the super-enhancer peaks are labeled. (B) Mean ATAC-seq read coverage, calculated by averaging all reads per million per base pair (RPM/bp) in the 50kb windows centered on all super-enhancers or all non-super-enhancers. (C) Gene expression of super-enhancer associated genes, other enhancer associated genes, and all genes in CHO-K1. Whiskers extend 1.5 times the inter-quartile range, and "x" defines the mean. (D) Circular genome view plotting CHO-K1 and 2C10 super-enhancers within all C. griseus scaffolds greater than 100kb. Each numbered box represents a scaffold and each point represents a super-enhancer.

The closest gene, including overlapping genes, were assigned to be associated with a super-enhancer. All other stitched peaks were considered as "other enhancers" and associated genes were assigned in the same way. Transcript abundance levels of genes associated with super-enhancers were higher than those of non-super-enhancer associated

genes or all genes (Figure 5-1C). Several functional classes were enriched among those super-enhancer associated genes, including signal transduction and cancer-related KEGG pathways (Table 5-1). Notable among those in the TNF signaling pathway were the upstream receptor *Tnfrsf1b* and downstream transcription factors *Creb1*, *Creb3l1*, and *Atf6b*. Similarly, in the PI3K-AKT signaling pathway, upstream receptors such as *Igfr1*, *Itgav*, and *Itga7*, signal transducers such as *Pik3ca*, and downstream cell cycle effectors such as *Cdk6* and *Myc* were represented in the super-enhancer-associated gene set.

Table 5-1: Selected significantly enriched KEGG pathways among super-enhancer associated genes in CHO-K1.

KEGG Pathway	**p−value**
TNF signaling pathway	1.0e−03
PI3K−Akt signaling pathway	1.1e−03
Proteoglycans in cancer	2.3e−03
Pathways in cancer	2.8e−03
Insulin resistance	4.5e−03
Endocytosis	8.6e−03

5.4.2 Stability of super-enhancers through genetic lineages

Similar to the CHO-K1 super-enhancers, the 2C10 super-enhancers had higher mean read coverage than other stitched peaks, while super-enhancer associated genes also had higher transcript levels than other enhancer associated genes (Figure 5-2). The normalized ATAC-signal of super-enhancers that were identified for CHO-K1 and 2C10 were arranged in circular genome view (Figure 5-1D). All super-enhancer regions were represented as circles among all scaffolds greater than 100kb in length, with the size of the circle corresponding to the relative ATAC signal of that super-enhancer. A large proportion

of CHO-K1 and 2C10 super-enhancers localized to the same genomic regions, with similarly large circles coinciding in the same regions.

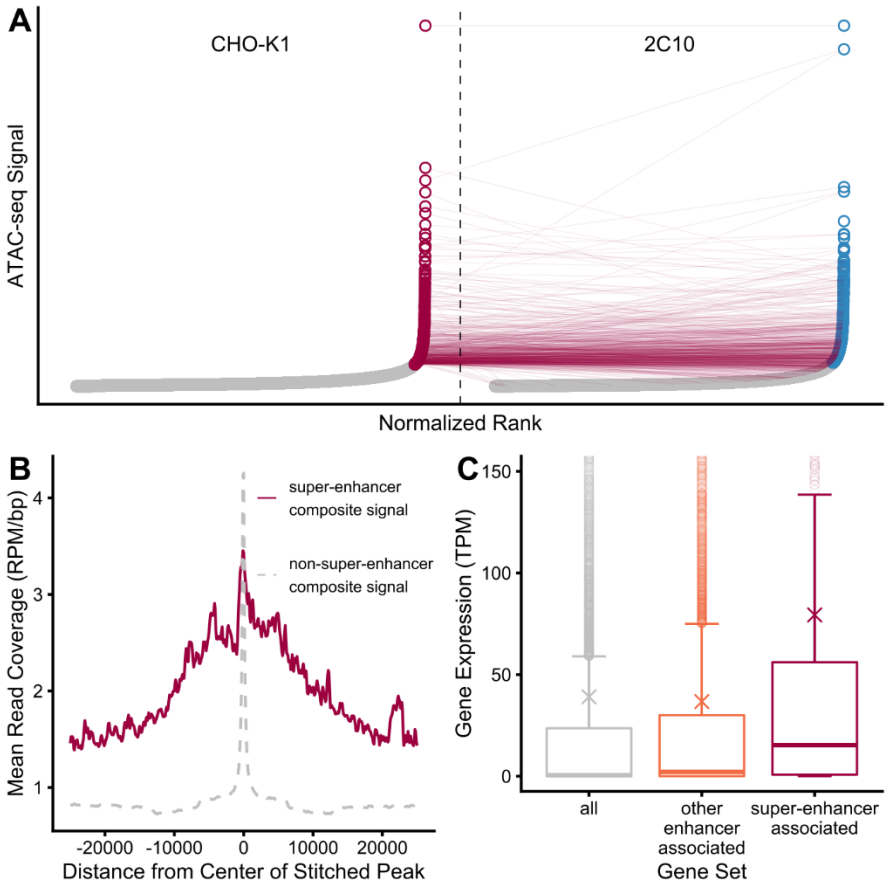

Figure 5-2: Super-enhancers in 2C10 and comparison to CHO-K1. (A) Super-enhancers from CHO-K1 mapped to 2C10 stitched peaks. Of 731 CHO-K1 super-enhancers, 72.6% (531) are also called super-enhancers in 2C10. Furthermore, 90% of CHO-K1 super-enhancers (658/731) are within the top 7.5% of 2C10 stitched peaks. (B) Mean ATAC-seq read coverage in the 50kb windows centered on all 2C10 super-enhancers or non-super-enhancers. (C) Gene expression of super-enhancer associated genes, other enhancer associated genes, and all genes in 2C10. Whiskers extend 1.5 times the inter-quartile range, and "x" defines the mean.

The majority of CHO-K1 super-enhancers (72.6%) were also identified as super-enhancers in 2C10 (Figure 5-2). Those which were not identified as super-enhancers in 2C10 fell mostly in the region immediately below super-enhancers (Figure 5-2). This finding suggested the stability of super-enhancers through cell lineages.

5.4.3 Inaccessible regions reveal genomic organization

In addition to identifying super-enhancers and peaks in active chromatin, ATAC-seq data also revealed regions of chromatin which were less accessible and likely not active. Regions between sequential peaks that were more than 100kb apart were considered inaccessible regions. An example of inaccessibility was the long stretch of genome (2.5Mbp) surrounding the *Dcc* gene that was devoid of peaks between *Mbd2* and *Mex3c* (Figure 5-3A). The CHO-K1 RNA-seq coverage shown in Figure 2A demonstrated the transcriptional activity of the flanking genes *Mbd2* and *Mex3c*. However, the *Dcc* and gene, which encodes a netrin 1 receptor for axon guidance of neurons and tumor suppression had very low transcriptional activity.

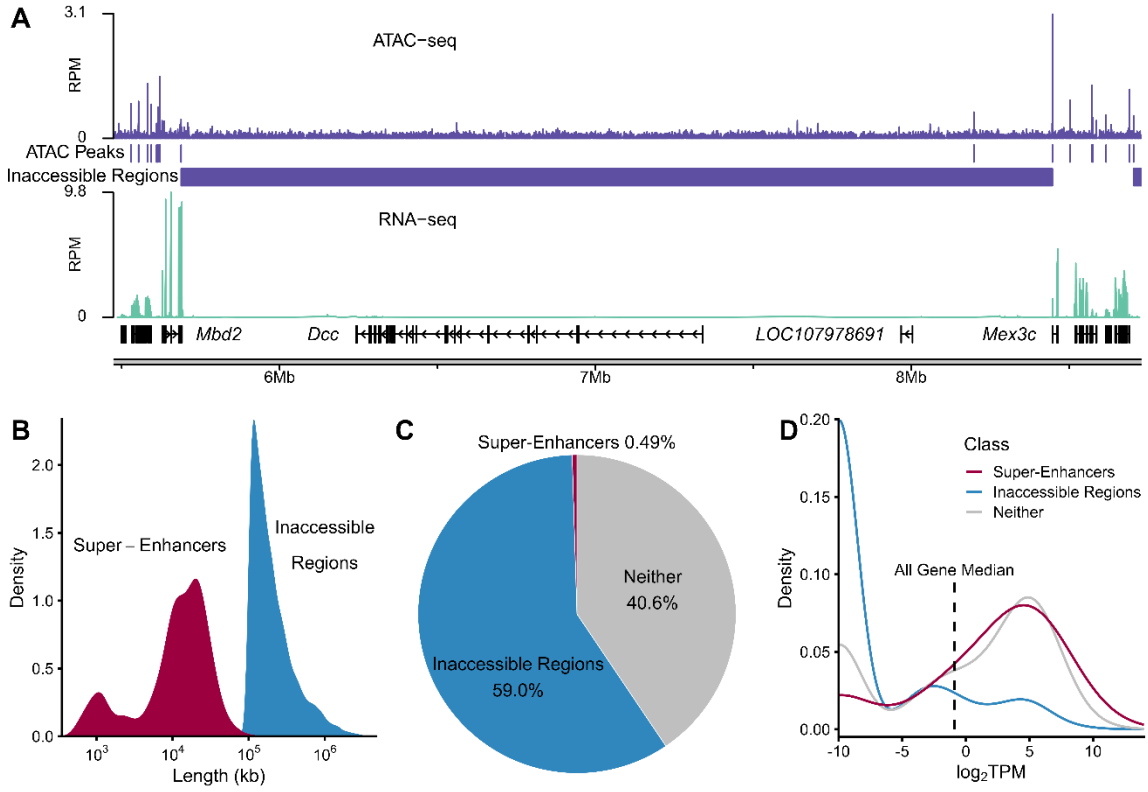

Figure 5-3: Inaccessible Regions in CHO-K1 and their associated genes.(A) Local genome view of representative inaccessible regions near the Dcc gene. (B) Length distributions of super-enhancers and inaccessible regions. (C) Proportion of the CHO genome covered by

super-enhancers and inaccessible regions. (D) Density plot of expression for genes associated with super-enhancers, inaccessible regions, or neither.

The length of super-enhancers spanned over 500-fold but rarely extended beyond 100kb (Figure 5-3B). In contrast, some inaccessible regions extended upwards of 1Mb. A large proportion of inaccessible regions were near 100kb. The choice of length criterion for calling inaccessibility thus affected the number of inaccessible regions in the genome. Using 100kb as the criterion, the inaccessible regions covered more than half of the genome, while super-enhancers represented a select subset of the genome (Figure 5-3C). If the criterion was set at 300kb, the proportion of the genome that was inaccessible was still over 30% (data not shown). Inaccessible regions thus constitute a very large fraction of the entire genome.

The genes in the inaccessible regions were significantly enriched for tissue-specific KEGG pathways such as sensory input and neuronal signaling (Table 5-2). Among the 1175 genes annotated in the KEGG olfactory transduction pathway, 914 were identified to be within inaccessible regions. Within the neuroactive ligand-receptor interaction pathway, inaccessible regions represented genes such as GABA receptor subunits *Gabra4* through *Gabra6,* histamine receptors *Hrh1* and *Hrh4,* and dopamine receptor *Drd2,* all of which had 0 TPM according to RNA-seq analysis.

Table 5-2: Selected significantly enriched KEGG pathways among genes in CHO-K1 inaccessible regions.

KEGG Pathway	**p−value**
Olfactory transduction	< 2.3e−308
Type I diabetes mellitus	1.3e−05
Graft−versus−host disease	2.9e−05
Neuroactive ligand−receptor interaction	3.1e−05

Taste transduction	2.5e−04
Retinol metabolism	5.6e−04

5.4.4 Impact of accessibility on the transcriptome

Since mRNA abundance is a balance between transcription and RNA turnover rate, the transcript level may not directly reflect the transcription rate. Nevertheless, the majority (74%) of super-enhancer associated genes were expressed above the median expression of all genes (Figure 5-3D). Compared to genes which were neither associated with super-enhancers nor inaccessible regions, super-enhancer associated genes had lower frequency in the non-expressed region and higher frequency in the highly expressed region. In contrast, the majority of genes within inaccessible regions had transcript levels below one tenth of the median level of all genes (0.5 TPM). However, 15% had expression levels higher than the median of all genes, while a small portion were expressed at high levels (> 5 TPM), including *Ghitm* and *Lta4h*. If regions with low RNA-seq read density or length greater than 300kb were used as the criterion, the resulting inaccessible regions had much fewer transcriptionally active genes (Figure 5-4). However, this also failed to identify many inactive regions correctly identified using 100kb as the criterion, decreasing method sensitivity.

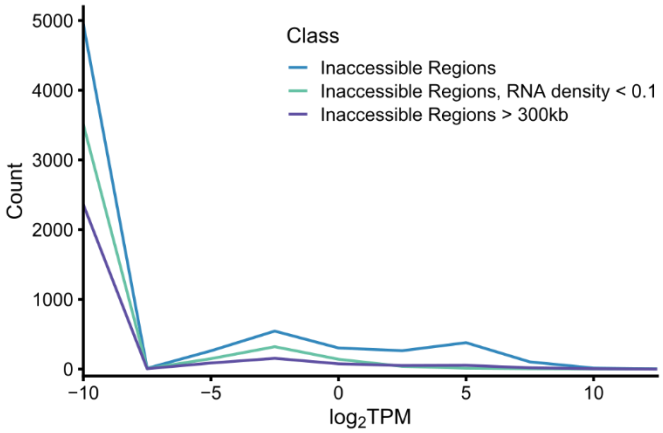

Figure 5-4: The number of genes associated with inaccessible regions. Binning was performed by transcript expression level as measured by RNA-seq TPM. When the originally defined inaccessible regions are filtered for RNA-seq read density in the region less than 0.1 times the genome-wide average or length greater than 300kb, fewer expressed genes are found but also fewer non-expressed genes are found.

5.4.5 Accessibility of Integration Sites of Transgene

In a previous study (O'Brien et al., 2018), a reporter gene (GFP) in a lentivirus vector was integrated at MOI of <=0.02 to isolate low-copy integrants. Through FACS the bottom 10% and top 1% GFP-expressing cells were pooled and integration sites were identified using PacBio sequencing. In addition, 34 top 1% single-copy clones were isolated and integration sites assessed by traditional Sanger sequencing of vector-genome junctions. The accessibilities of their integration sites were evaluated. Of these 34 clones, four had their integration sites in or within 50kb of a super-enhancer (Figure 5-5A). The integration sites of the top 1% pool showed a marginally higher representation near super-enhancers than the bottom 10% pool and a random sampling of 10,000 genomic regions, while the top 1% single-copy clones showed significant enrichment for super-enhancer regions compared to those populations (Table 5-3).

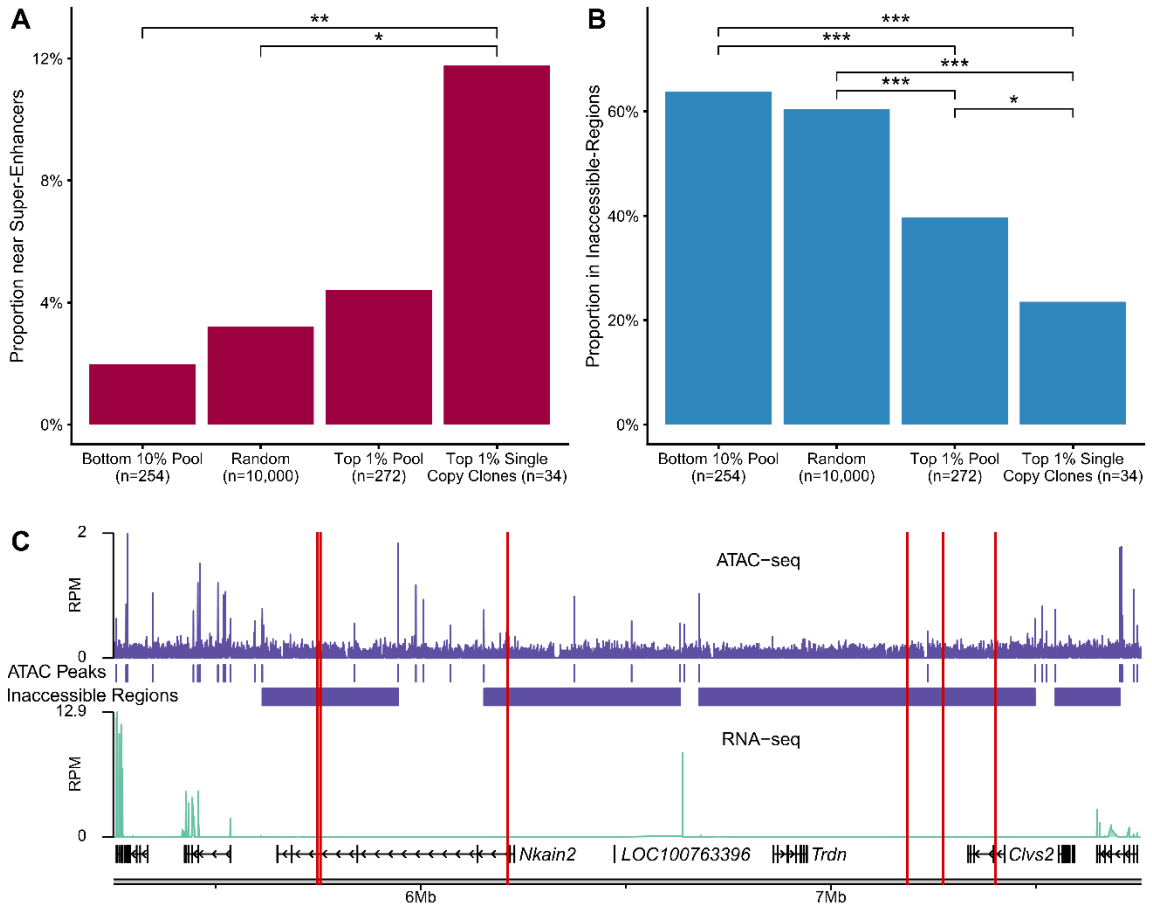

Figure 5-5: Intersecting integration site data with epigenomic features. Integration sites of randomly integrated GFP transgene tested for whether they are (A) within 50kb of a super-enhancer or (B) overlap an inaccessible region. Frequencies were compared pairwise using the 2x2 Boschloo's one-tailed exact test. (C) A local genome view of a region which had six distinct integration events (red lines) among the bottom 10% population. *p<0.05, **p<0.01, ***p<1e-5.

Table 5-3: Significance of pair-wise comparisons of lentiviral integration sites within 50kb of a super-enhancer. The p-values calculated via pair-wise Boschloo one-tailed exact tests. *p<0.05, **p<0.01.

	Bottom10%	*Random*	*Top1%*	*Top1%Clones*
Bottom10%		0.16	0.079	0.0086**
Random			0.16	0.018*
Top1%				0.065
Top1%Clones				

Only eight out of the 34 top 1% single-copy clones were found to be associated with inaccessible regions (Figure 5-5B). One of the eight had a very low regional RNA-seq read coverage (< 0.1 times the genome-wide average) near the integration site and had the lowest level of GFP transcripts among all clones, making it most likely a mis-sorted clone in FACS operation. In the remaining seven top 1% clone integration sites, the associated inaccessible regions overlapped transcriptionally active genes. Most of these genes were extremely long (> 95kb) with an ATAC peak at the transcription start site, but with long intragenic regions without peaks. This partially explained why these were identified as inaccessible regions. In comparison, over half of the integration sites of the bottom 10% pool were in inaccessible regions, a level comparable to that of random sites and significantly higher than that of the top 1% pool or the top 1% single-copy clones (Figure 5-5B and Table 5-4).

Table 5-4: Significance of pair-wise comparisons of lentiviral integration sites in inaccessible regions. The p-values calculated via pair-wise Boschloo one-tailed exact tests. *p<0.05, **p<0.01, ***p<1e-5.

	Bottom10%	*Random*	*Top1%*	*Top1%Clones*
Bottom10%		1.4e−01	1.8e−08***	4.2e−06***
Random			5.4e−12***	9.1e−06***
Top1%				3.3e−02*
Top1%Clones				

Interestingly, a 2Mb stretch of inaccessible regions harbored six distinct integration events from the bottom 10% pool, suggesting that certain genomic locations were especially poor sites for exogenous gene expression (Figure 5-5C). No top 1% pool or clone integration events were detected in this region.

5.4.6 Site-directed integration into one super-enhancer

The data presented above indicated that product genes integrated near super-enhancer regions were more likely to have high transcript levels. Thus, the ideal candidate for site-directed gene integration should be a super-enhancer in the top tier of transcriptional activity (Figure 5-6A). Additionally, it should be a super-enhancer in both CHO-K1 and 2C10, and have no evidence of genomic instability as assayed by comparative genomic hybridization (Bandyopadhyay et al., 2019). We chose a super-enhancer region that met these criteria and was among the top 1% GFP-expressing single copy clones' integration sites. The region included the *Plec* gene (Figure 5-6B). An IgG expression construct was built for CRISPR-mediated integration with destabilized GFP as a positive selection marker as described previously (O'Brien et al., 2018). To reduce capturing random integration events, BFP and HSVtk were included in the backbone as negative selection markers (see Table 5-5). Integration was targeted to two loci: one proximal to the super-enhancer (sites A and B) in the same vicinity as a single-copy clone, the other 50kb from the opposite end of the super-enhancer (sites C and D). In each locus, the transgene cassette was integrated in both directions.

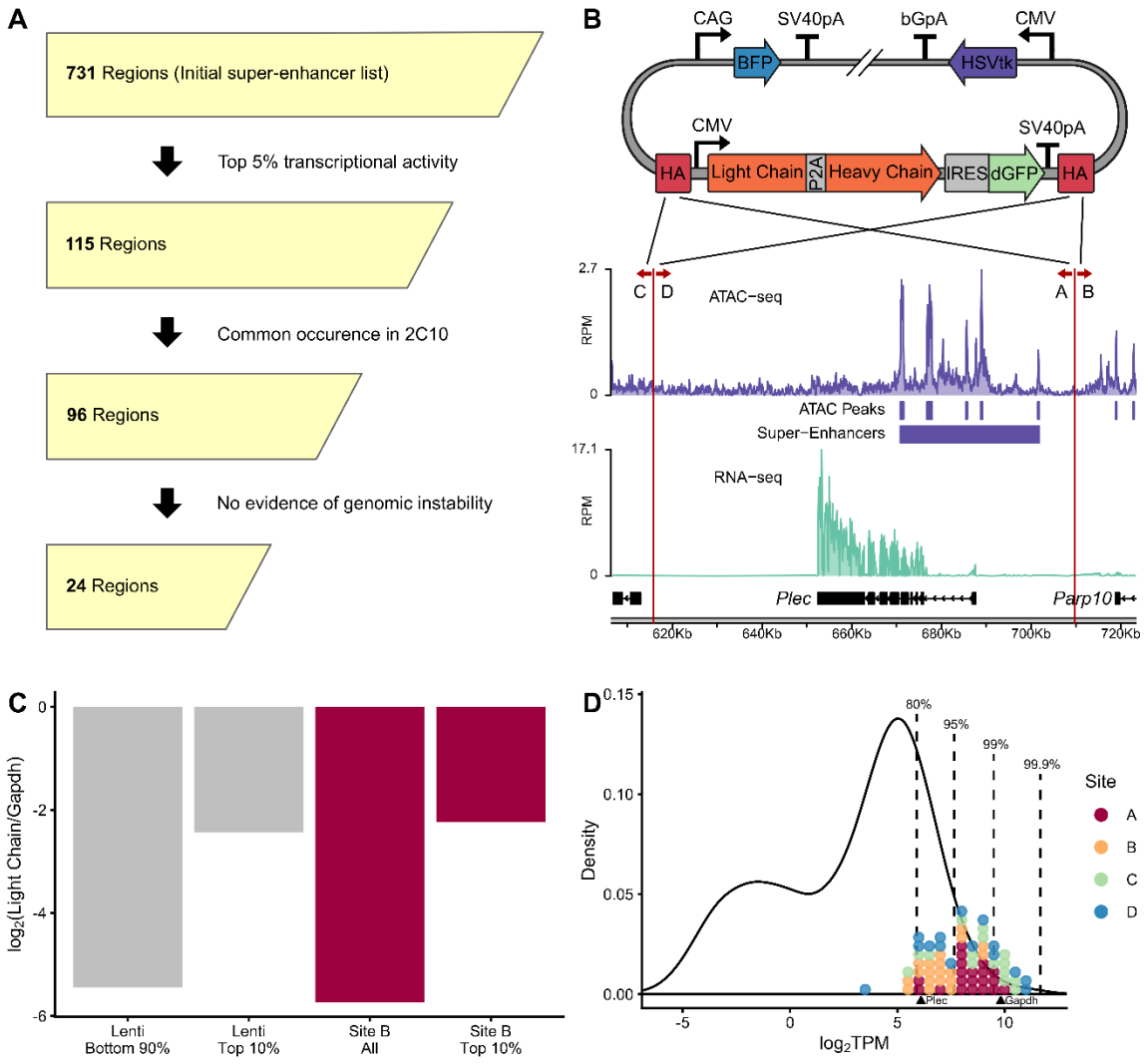

Figure 5-6: CRISPR-mediated integration into putative super-enhancer site. (A) The approach taken to select the super-enhancer for testing. (B) The gene construct and integration orientations around the *Plec* super-enhancer site. The "B" orientation matches one single-copy GFP clone identified in the lentiviral integration screen. (C) The transcript abundance of IgG light chain in the lentiviral integration sub-pools and the site-specific integration sub-pools. (D) The overall RNA-seq distribution of non-zero-TPM genes in the CHO-K1 transcriptome. The qRT-PCR expression level of all isolated clones was converted to TPM using Gapdh as a common reference.

Table 5-5: Clone selection statistics. The number of isolated clones from fluorescence-activated cell sorting and subsequent number of clones which passed further selection and screening during integration into the *Plec* super-enhancer region.

	Number of Clones	% of Parent
BFP-negative, isolated	186	NA
GCV-resistant	136	73
Assayed by Junction PCR	92	68
Correct Junctions	69	75

Upon CRISPR-mediated transgene integration, GFP-expressing cells were FACS-sorted into single-cell clones or pools. The GFP-expressing cell pool from integration site B and the top 10% subset of that pool were assayed for IgG light chain transcript level by qRT-PCR (Figure 5-6C). The light and heavy chain transcripts from this expression cassette were previously shown to be similar (O'Brien et al., 2018). The light chain transcript level was also correlated to GFP fluorescence intensity (Figure 5-7). Transcript levels of the top 10% GFP-expressing cell pool was comparable to the top 10% pool obtained via lentiviral random integration described above. The IgG light chain transcript levels in the 69 clones across the four integration sites (denoted by color) were denoted on the transcript abundance distribution curve (Figure 5-6D). Except for seven clones, all were in the top 20% abundance level of expressed cellular transcripts. Most clones also possessed higher light chain transcript abundance than *Plec*, the gene overlapping the super-enhancer. More than half were in the top 5% of expressed transcripts. As a reference, the expression level of *Gapdh*, one of the most abundant genes in CHO cells and a commonly used reference gene, was also marked. Most clones expressed light chain within a range between 2-fold higher to 8-fold lower than *Gapdh* (Figure 5-6D), demonstrating that integrating near a super-enhancer can yield high transcription of transgene. However, the resulting clones still exhibited transcript levels spanning one order of magnitude in range. GFP expression of five representative clones was assessed after 50 generations of

culture and persistent high GFP expression was seen in all clones. The number of clones examined was small, but nevertheless the apparent stability was consistent with the finding that super-enhancers were stable through cell lineages (Figure 5-2).

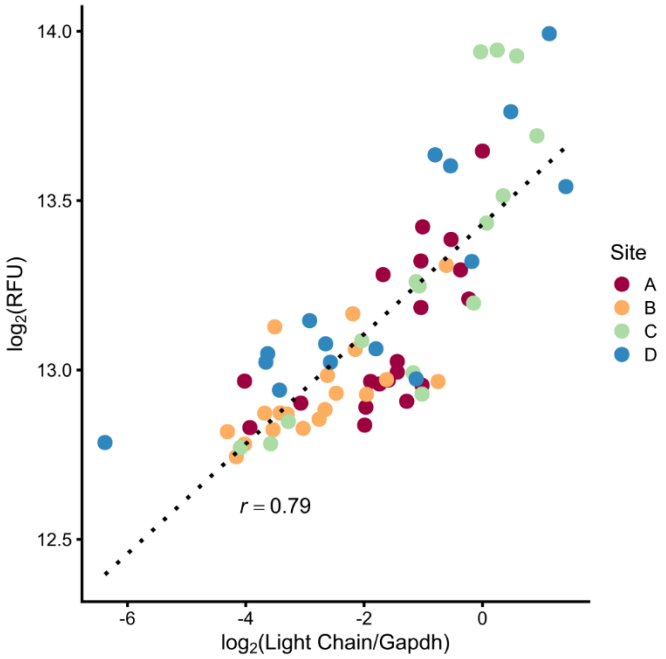

Figure 5-7: The relationship between GFP fluorescence to IgG light chain transcript level in targeted integration clones. Pearson's correlation value is represented by r. The GFP relative fluorescence units (RFU) measurements were taken using a BioTek Synergy H1 microplate reader, and the light chain transcript level was assessed by qRT-PCR.

5.5 Discussion

5.5.1 Super-enhancers and Inaccessible Regions Influence Cell Identity

In this study we used ATAC-seq to identify super-enhancers comprised of clusters of highly accessible open chromatin peaks in CHO cells. Most super-enhancer-associated genes had higher transcript levels than the median of all genes with functional enrichment in signal transduction and cancer-related pathways. This may hint that the immortalization of CHO cells may involve super-enhancer-mediated upregulation of oncogenes in addition to the reported mutations in p53 (Vishwanathan et al., 2015) and amplification of cMyc

(Vishwanathan et al., 2017). Similar super-enhancer interactions with oncogenes have been reported in cancers such as multiple myeloma (Loven et al., 2013), triple-negative breast cancer (Wang et al., 2015), and medulloblastoma (Lin et al., 2016), among others.

Comparison of CHO-K1 and a derived IgG-producing cell line 2C10 showed that super-enhancers can be stably inherited for a large number of cell divisions. A recent computational study analyzed 200 cell types to identify the core transcription factors which establish cell-type identity (D'Alessio et al., 2015). Among the top ten core transcription factors attributed to human immortalized ovary epithelium tissues, eight were mappable to the CHO genome, and three (*E2f7*, *Ets1*, *Fosl1*) were proximal to super-enhancers in CHO-K1. Another two (*Stat1*, *Hmga1*) were found among the top 8% of CHO-K1 stitched peaks, immediately below the inflection point. All of these transcription factors are involved in regulation of cell proliferation and likely to be important for maintaining the immortalized phenotype. Super-enhancer analysis was also performed on Chinese hamster induced pluripotent stem cells (Pei et al., 2017) and the *Sox2* and *Klf4* pluripotency genes were found to be associated with super-enhancers (data not shown). This supports the idea that super-enhancers contribute to maintaining specific cell phenotypes and are epigenetically stable in a cell line.

In contrast to super-enhancers, inaccessible regions with long stretches of genome without any ATAC-seq peaks were rich in silent genes specific to tissues such as the neuronal and sensory organs. These results were consistent with previous reports which found genes without accessible promoters were enriched for tissue-specific genes such as rhodopsin-like receptors and olfactory receptors (Boyle et al., 2008). Important to establishing the identity of a cell type is not only the expression of genes associated with

the cell type, but also the suppression of genes specific to other cell types through inaccessibility of chromatin. These regions were also found to be stably inherited through many cell divisions as 97% of inaccessible regions in the 2C10 cell line overlapped inaccessible regions in CHO-K1 (data not shown).

Before ATAC became commonly used, studies aimed at annotating the epigenome employed ChIP-seq or DNase-seq to cluster genome regions into distinct chromatin states. One study classified large proportions of the genome as repressed, with low gene expression levels and higher frequencies of transgene silencing when a gene is inserted into these regions (Filion et al., 2010). Another study described quiescent-state chromatin as devoid of transcription and transcription factor binding (Hoffman et al., 2013). Both studies also reported these inactive regions to be longer than active areas. Here, we used a single assay, ATAC-seq, to identify inaccessible regions which shared similar characteristics with the described repressed and quiescent chromatin states. Given that only about half of the genes encoded in a mammalian genome are expressed in a given cell type, it is not surprising that at least half of the genome in CHO would be inaccessible.

5.5.2 Integration site accessibility and GFP expression

Random integration of a GFP- lentiviral vector revealed that the top 1% of GFP expressing clones had significantly higher frequencies of integration into super-enhancer regions. Though super-enhancers constituted only about 0.49% of the entire genome, 12% of the top 1% clones integrated into super-enhancer regions. It suggests a higher propensity of these regions to promote high transcriptional activity. The top 1% single-copy clones showed even greater enrichment for super-enhancers than the top 1% pool likely because the pool harbored “passenger” integration events in addition to integration in super-

enhancer regions. During top 1% clone screening, 46% of clones were found to be multiple-copy, reflecting the frequency of multiple-copy integrations in the top 1% pool. About 60% of random integrants as well as those in the bottom 10% of GFP intensity had the vector integrated into an inaccessible region, significantly higher than those in the top 1%. The results indicate that integration into accessible or inaccessible regions affects the chance of expression but does not guarantee high or low reporter activity.

5.5.3 Comparison of super-enhancers to hot-spots and safe-harbors

There has been a heightened interest in targeting integration into a specific genomic locus for transgene expression. The desired target site should promote a high level of transcription and confer long-term structural and gene expression stability. One study identified the *Fer1L4* gene as a "hot-spot" in the genome through screening and targeted a product gene to the site (Zhang et al., 2015). Multiple studies targeted transgene integration into the glycosylation chaperone *COSMC* gene as a "safe-harbor" site for consistent and stable expression (J. S. Lee, Kallehauge, Pedersen, & Kildegaard, 2015; Ronda et al., 2014). These two sites were not among the CHO-K1 super-enhancers identified in this study, nor were they in inaccessible regions. Yet both sites were in genomic neighborhoods with high density of ATAC peaks and transcriptionally active genes. This further demonstrates that being in a super enhancer site increases the probability of high transcriptional activity, but it is not a necessary condition.

Integration into the *Plec* super-enhancer resulted in over half of clones expressing the transgene at high levels—in the top 5% of all expressed cellular genes (Figure 5-3D). But even with a strong CMV promoter the expression level was not universally higher than that of a highly abundant housekeeping gene, *Gapdh*. In *Drosophila* some clusters of

multiple enhancers were ascribed to ensure a gene was expressed, but not necessarily to a very high level of expression (Perry, Boettiger, & Levine, 2011). This is consistent with the results shown in Figure 5-3D, where most genes are expressed at high but not necessarily at the highest levels. It has been reported that integration of a transgene into a predisposed transcriptionally active site identified in a high producer was highly active but could still be further enhanced by integrating two copies of the transgene into the same site (J. Carver et al., 2020).

5.5.4 Variability of transcript level in clones integrated in super-enhancer regions

It was reported that CRISPR-mediated integration of mCherry into CHO cells led to a more narrow range of fluorescence than clones created through random integration (J. S. Lee et al., 2015). In contrast, another study showed targeted integration of IgG resulted in clones with variable expression levels (J. Carver et al., 2020). Although most clones integrated near the *Plec* super-enhancer locus had high transcript levels of IgG, the expression level did span over a range. A number of factors could have caused the variability in the transcript levels. There are likely two alleles of the *Plec* gene in the CHO genome that may harbor some differences. The CMV promoter is often methylated which affects its transcriptional activity (M. Kim, O'Callaghan, Droms, & James, 2011; Moritz, Becker, & Gopfert, 2015). The CHO genome is inherently susceptible to karyotypic changes (Bandyopadhyay et al., 2019) resulting in a heterogeneous population that may also contribute to transcriptional activity variation.

Despite integration in the vicinity of the *Plec* super-enhancer, the IgG titer measured during clonal expansion showed a range of distribution among the clones; some were similar to or even higher than the producing clone 2C10, while others were lower

(Figure 5-8). This titer distribution reminds us that the overall productivity is determined by a myriad of factors. A high level of transcription is just one component of a high producer cell line. Other traits are also important, including translation, protein folding, mitochondrial activity and biogenesis (Nissom et al., 2006), metabolism, protein secretion, redox balance, and growth/death control (Seth, Charaniya, Wlaschin, & Hu, 2007). Integration to a super-enhancer site facilitates accomplishing the critical role of high transgene transcript level. It is still beneficial to screen the resulting clones to isolate the highest producers which possess the additional traits of hyper-productivity.

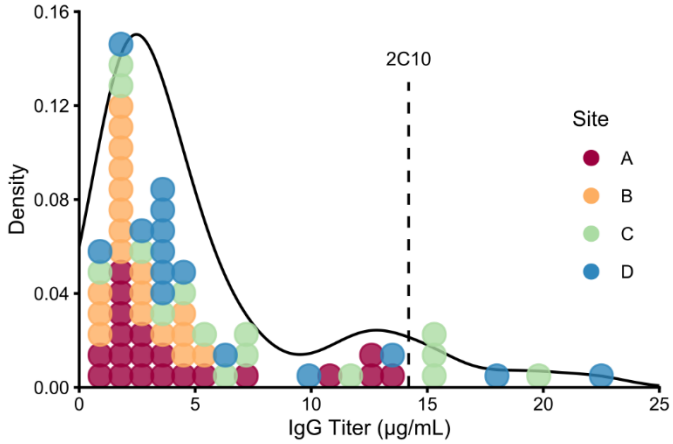

Figure 5-8: The distribution of IgG titer among *Plec*-targeted integration clones. Cells were split at a 1:6 ratio into a 24-well plate with 0.5 mL media. Supernatant was harvested on day three. Supernatant was diluted 1:400 and assayed by ELISA. The dotted vertical line represents the average titer of 2C10 assayed the same way. Several clones showed similar or greater titer to 2C10, a top 1% FACS-sorted cell line reported to have a specific productivity of 18 pg/cell/day (O'Brien et al., 2018).

5.5.5 Concluding remarks

Using ATAC-seq to identify super-enhancers and inaccessible regions in the CHO genome can facilitate the design or screening of cell lines for transgene expression. When random integration is used, the integration site can be identified by integration site analysis (O'Brien, Ojha, Wu, & Hu, 2020), and then ATAC-determined accessibility information

can be used in the selection of a production clone. Random integration methods have a high likelihood of inserting transgenes into inaccessible regions, which cover over half of the genome, leading to poor expression. Simply avoiding inaccessible regions may greatly narrow down the pool of candidate high-producers. Further, transgene integration into a super-enhancer region increases the probability of having high transcript expression. Since super-enhancers are epigenetically stable through cell lineages, integration into a super-enhancer region can possibly confer long-term gene expression stability. Simultaneous targeting of two or three super-enhancer loci may offer the best chance at obtaining stable high producers.

6 Bibliography

Balasubramanian, S., Wurm, F. M., & Hacker, D. L. (2016). Multigene expression in stable CHO cell pools generated with the piggyBac transposon system. *Biotechnology Progress, 32*(5), 1308-1317. doi:https://doi.org/10.1002/btpr.2319

Banaszynski, L. A., Chen, L.-C., Maynard-Smith, L. A., Ooi, A. G. L., & Wandless, T. J. (2006). A rapid, reversible, and tunable method to regulate protein function in living cells using synthetic small molecules. *Cell, 126*(5), 995-1004. doi:10.1016/j.cell.2006.07.025

Bandyopadhyay, A. A., O'Brien, S. A., Zhao, L., Fu, H. Y., Vishwanathan, N., & Hu, W. S. (2019). Recurring genomic structural variation leads to clonal instability and loss of productivity. *Biotechnol Bioeng, 116*(1), 41-53. doi:10.1002/bit.26823

Barnes, L. M., Bentley, C. M., & Dickson, A. J. (2003). Stability of protein production from recombinant mammalian cells. *Biotechnol Bioeng, 81*(6), 631-639. doi:10.1002/bit.10517

Boyle, A. P., Davis, S., Shulha, H. P., Meltzer, P., Margulies, E. H., Weng, Z., . . . Crawford, G. E. (2008). High-resolution mapping and characterization of open chromatin across the genome. *Cell, 132*(2), 311-322. doi:10.1016/j.cell.2007.12.014

Brough, D. E., Lizonova, A., Hsu, C., Kulesa, V. A., & Kovesdi, I. (1996). A gene transfer vector-cell line system for complete functional complementation of adenovirus early regions E1 and E4. *Journal of virology, 70*(9), 6497-6501. doi:10.1128/jvi.70.9.6497-6501.1996

Buenrostro, J. D., Wu, B., Chang, H. Y., & Greenleaf, W. J. (2015). ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Curr Protoc Mol Biol, 109*, 21 29 21-21 29 29. doi:10.1002/0471142727.mb2129s109

Cadena-Herrera, D., Esparza-De Lara, J. E., Ramírez-Ibañez, N. D., López-Morales, C. A., Pérez, N. O., Flores-Ortiz, L. F., & Medina-Rivero, E. (2015). Validation of three viable-cell counting methods: Manual, semi-automated, and automated. *Biotechnology Reports, 7*, 9-16. doi:https://doi.org/10.1016/j.btre.2015.04.004

Carver, J., Ng, D., Zhou, M., Ko, P., Zhan, D., Yim, M., . . . Hu, Z. (2020). Maximizing antibody production in a targeted integration host by optimization of subunit gene dosage and position. *Biotechnology Progress, 36*(4), e2967. doi:https://doi.org/10.1002/btpr.2967

Carver, J., Ng, D., Zhou, M., Ko, P., Zhan, D., Yim, M., . . . Hu, Z. (2020). Maximizing antibody production in a targeted integration host by optimization of subunit gene dosage and position. *Biotechnol Prog*, e2967. doi:10.1002/btpr.2967

Chadeuf, G., Favre, D., Tessier, J., Provost, N., Nony, P., Kleinschmidt, J., . . . Salvetti, A. (2000). Efficient recombinant adeno-associated virus production by a stable rep-cap HeLa cell line correlates with adenovirus-induced amplification of the integrated rep-cap genome. *The Journal of Gene Medicine, 2*(4), 260-268. doi:https://doi.org/10.1002/1521-2254(200007/08)2:4<260::AID-JGM111>3.0.CO;2-8

Chejanovsky, N., & Carter, B. J. (1989). Mutagenesis of an AUG codon in the adeno-associated virus rep gene: Effects on viral DNA replication. *Virology, 173*(1), 120-

128. doi:https://doi.org/10.1016/0042-6822(89)90227-4
Chen, C.-L., Clark, R., Chen, R., Schnepp, B. C., & Johnson, P. R. (2002). Generation of Stable 293 Cell Lines for the Production of rAAV Vectors. *Molecular Therapy, 5*(5), S47. doi:10.1016/S1525-0016(16)42967-9
Clark, K. R., Voulgaropoulou, F., Fraley, D. M., & Johnson, P. R. (1995). Cell Lines for the Production of Recombinant Adeno-Associated Virus. *Human Gene Therapy, 6*(10), 1329-1341. doi:10.1089/hum.1995.6.10-1329
Clément, N., & Grieger, J. C. (2016). Manufacturing of recombinant adeno-associated viral vectors for clinical trials. *Molecular Therapy - Methods & Clinical Development, 3*, 16002. doi:https://doi.org/10.1038/mtm.2016.2
Clément, N., Knop, D. R., & Byrne, B. J. (2009). Large-Scale Adeno-Associated Viral Vector Production Using a Herpesvirus-Based System Enables Manufacturing for Clinical Studies. *Human Gene Therapy, 20*(8), 796-806. doi:10.1089/hum.2009.094
Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., . . . Zhang, F. (2013). Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science, 339*, 819-823.
Coronel, J., Patil, A., Al-Dali, A., Braβ, T., Faust, N., & Wissing, S. (2021). Efficient Production of rAAV in a Perfusion Bioreactor Using an ELEVECTA® Stable Producer Cell Line. *Genetic Engineering & Biotechnology News, 41*(S2), S23-S23. doi:10.1089/gen.41.S2.07
D'Alessio, A. C., Fan, Z. P., Wert, K. J., Baranov, P., Cohen, M. A., Saini, J. S., . . . Young, R. A. (2015). A Systematic Approach to Identify Candidate Transcription Factors that Control Cell Identity. *Stem Cell Reports, 5*(5), 763-775. doi:10.1016/j.stemcr.2015.09.016
Dobnik, D., Kogovšek, P., Jakomin, T., Košir, N., Tušek Žnidarič, M., Leskovec, M., . . . Ravnikar, M. (2019). Accurate Quantification and Characterization of Adeno-Associated Viral Vectors. *Frontiers in Microbiology, 10*, 1570. Retrieved from https://www.frontiersin.org/article/10.3389/fmicb.2019.01570
El Andari, J., & Grimm, D. (2021). Production, Processing, and Characterization of Synthetic AAV Gene Therapy Vectors. *Biotechnology Journal, 16*(1), 2000025. doi:https://doi.org/10.1002/biot.202000025
Farson, D., Harding, T. C., Tao, L., Liu, J., Powell, S., Vimal, V., . . . Donahue, B. A. (2004). Development and characterization of a cell line for large-scale, serum-free production of recombinant adeno-associated viral vectors. *The Journal of Gene Medicine, 6*(12), 1369-1381. doi:https://doi.org/10.1002/jgm.622
Ferrari, F. K., Samulski, T., Shenk, T., & Samulski, R. J. (1996). Second-strand synthesis is a rate-limiting step for efficient transduction by recombinant adeno-associated virus vectors. *Journal of virology, 70*(5), 3227-3234. doi:10.1128/jvi.70.5.3227-3234.1996
Filion, G. J., van Bemmel, J. G., Braunschweig, U., Talhout, W., Kind, J., Ward, L. D., . . . van Steensel, B. (2010). Systematic protein location mapping reveals five principal chromatin types in Drosophila cells. *Cell, 143*(2), 212-224. doi:10.1016/j.cell.2010.09.009
François, A., Bouzelha, M., Lecomte, E., Broucque, F., Penaud-Budloo, M., Adjali, O., . . .

Ayuso, E. (2018). Accurate Titration of Infectious AAV Particles Requires Measurement of Biologically Active Vector Genomes and Suitable Controls. *Molecular Therapy - Methods & Clinical Development, 10*, 223-236. doi:10.1016/j.omtm.2018.07.004

Gao, G.-P., Qu, G., Faust, L. Z., Engdahl, R. K., Xiao, W., Hughes, J. V., . . . Wilson, J. M. (1998). High-Titer Adeno-Associated Viral Vectors from a Rep/Cap Cell Line and Hybrid Shuttle Virus. *Human Gene Therapy, 9*(16), 2353-2362. doi:10.1089/hum.1998.9.16-2353

Gonçalves, M. A. F. V. (2005). Adeno-associated virus: from defective virus to effective vector. *Virology Journal, 2*(1), 43. doi:10.1186/1743-422X-2-43

Grimm, D., Kern, A., Pawlita, M., Ferrari, F. K., Samulski, R. J., & Kleinschmidt, J. A. (1999). Titration of AAV-2 particles via a novel capsid ELISA: packaging of genomes can limit production of recombinant AAV-2. *Gene Therapy, 6*(7), 1322-1330. doi:10.1038/sj.gt.3300946

Grimm, D., Lee, J. S., Wang, L., Desai, T., Akache, B., Storm, T. A., & Kay, M. A. (2008). In Vitro and In Vivo Gene Therapy Vector Evolution via Multispecies Interbreeding and Retargeting of Adeno-Associated Viruses. *Journal of virology, 82*(12), 5887-5911. doi:10.1128/JVI.00254-08

He, X., Urip, B. A., Zhang, Z., Ngan, C. C., & Feng, B. (2021). Evolving AAV-delivered therapeutics towards ultimate cures. *Journal of Molecular Medicine, 99*(5), 593-617. doi:10.1007/s00109-020-02034-2

Hidalgo, P., Anzures, L., Hernández-Mendoza, A., Guerrero, A., Wood Christopher, D., Valdés, M., . . . Sandri-Goldin, R. M. (2016). Morphological, Biochemical, and Functional Study of Viral Replication Compartments Isolated from Adenovirus-Infected Cells. *Journal of virology, 90*(7), 3411-3427. doi:10.1128/JVI.00033-16

Hoffman, M. M., Ernst, J., Wilder, S. P., Kundaje, A., Harris, R. S., Libbrecht, M., . . . Noble, W. S. (2013). Integrative annotation of chromatin elements from ENCODE data. *Nucleic Acids Res, 41*(2), 827-841. doi:10.1093/nar/gks1284

Hölscher, C., Hörer, M., Kleinschmidt, J. A., Zentgraf, H., Bürkle, A., & Heilbronn, R. (1994). Cell lines inducibly expressing the adeno-associated virus (AAV) rep gene: requirements for productive replication of rep-negative AAV mutants. *Journal of virology, 68*(11), 7169-7177. doi:10.1128/JVI.68.11.7169-7177.1994

Huang, Y., Li, Y., Wang, Y. G., Gu, X., Wang, Y., & Shen, B. F. (2007). An efficient and targeted gene integration system for high-level antibody expression. *J Immunol Methods, 322*(1-2), 28-39. doi:10.1016/j.jim.2007.01.022

Inniss, M. C., Bandara, K., Jusiak, B., Lu, T. K., Weiss, R., Wroblewska, L., & Zhang, L. (2017). A novel Bxb1 integrase RMCE system for high fidelity site-specific integration of mAb expression cassette in CHO Cells. *Biotechnol Bioeng, 114*(8), 1837-1846. doi:10.1002/bit.26268

Kallunki, T., Barisic, M., Jäättelä, M., & Liu, B. (2019). How to Choose the Right Inducible Gene Expression System for Mammalian Studies? *Cells, 8*(8). doi:10.3390/cells8080796

Kärber, G. (1931). Beitrag zur kollektiven Behandlung pharmakologischer Reihenversuche. *Naunyn-Schmiedebergs Archiv für experimentelle Pathologie und Pharmakologie, 162*(4), 480-483. doi:10.1007/BF01863914

Kim, M., O'Callaghan, P. M., Droms, K. A., & James, D. C. (2011). A mechanistic understanding of production instability in CHO cell lines expressing recombinant monoclonal antibodies. *Biotechnol Bioeng, 108*(10), 2434-2446. doi:10.1002/bit.23189

Kim, S. J., & Lee, G. M. (1999). Cytogenetic analysis of chimeric antibody-producing CHO cells in the course of dihydrofolate reductase-mediated gene amplification and their stability in the absence of selective pressure. *Biotechnology and Bioengineering, 64*(6), 741-749. doi:10.1002/(sici)1097-0290(19990920)64:6<741::Aid-bit14>3.0.Co;2-x

King, J. A., Dubielzig, R., Grimm, D., & Kleinschmidt, J. A. (2001). DNA helicase-mediated packaging of adeno-associated virus type 2 genomes into preformed capsids. *The EMBO Journal, 20*(12), 3282-3291. doi:https://doi.org/10.1093/emboj/20.12.3282

Kingston, R. E., Kaufman, R. J., Bebbington, C. R., & Rolfe, M. R. (2002). Amplification Using CHO Cell Expression Vectors. *Current Protocols in Molecular Biology, 60*(1), 16.23.11-16.23.13. doi:10.1002/0471142727.mb1623s60

Kitada, T., DiAndreth, B., Teague, B., & Weiss, R. (2018). Programming gene and engineered-cell therapies with synthetic biology. *Science, 359*(6376), eaad1067. doi:10.1126/science.aad1067

Kito, M., Itami, S., Fukano, Y., Yamana, K., & Shibui, T. (2002). Construction of engineered CHO strains for high-level production of recombinant proteins. *Appl Microbiol Biotechnol, 60*(4), 442-448. doi:10.1007/s00253-002-1134-1

Kohlbrenner, E., Aslanidi, G., Nash, K., Shklyaev, S., Campbell-Thompson, M., Byrne, B. J., . . . Zolotukhin, S. (2005). Successful Production of Pseudotyped rAAV Vectors Using a Modified Baculovirus Expression System. *Molecular Therapy, 12*(6), 1217-1225. doi:https://doi.org/10.1016/j.ymthe.2005.08.018

Kovesdi, I., & Hedley, S. J. (2010). Adenoviral Producer Cells. *Viruses, 2*(8). doi:10.3390/v2081681

Lee, J. S., Kallehauge, T. B., Pedersen, L. E., & Kildegaard, H. F. (2015). Site-specific integration in CHO cells mediated by CRISPR/Cas9 and homology-directed DNA repair pathway. *Sci Rep, 5*, 8572. doi:10.1038/srep08572

Lee, J. W., Devanarayan, V., Barrett, Y. C., Weiner, R., Allinson, J., Fountain, S., . . . Wagner, J. A. (2006). Fit-for-Purpose Method Development and Validation for Successful Biomarker Measurement. *Pharmaceutical Research, 23*(2), 312-328. doi:10.1007/s11095-005-9045-3

Lee, Z., Raabe, M., & Hu, W.-S. (2021). Epigenomic features revealed by ATAC-seq impact transgene expression in CHO cells. *Biotechnology and Bioengineering, 118*(5), 1851-1861. doi:https://doi.org/10.1002/bit.27701

Leppard, K. N. (1997). E4 gene function in adenovirus, adenovirus vector and adeno-associated virus infections. *Journal of General Virology, 78*(9), 2131-2138. doi:https://doi.org/10.1099/0022-1317-78-9-2131

Li, J., Samulski, R. J., & Xiao, X. (1997). Role for highly regulated rep gene expression in adeno-associated virus vector production. *Journal of virology, 71*(7), 5236-5243. doi:10.1128/jvi.71.7.5236-5243.1997

Lin, C. Y., Erkek, S., Tong, Y., Yin, L., Federation, A. J., Zapatka, M., . . . Northcott, P.

A. (2016). Active medulloblastoma enhancers reveal subgroup-specific cellular origins. *Nature, 530*(7588), 57-62. doi:10.1038/nature16546

Lock, M., McGorray, S., Auricchio, A., Ayuso, E., Beecham, E. J., Blouin-Tavel, V., . . . Snyder, R. O. (2010). Characterization of a Recombinant Adeno-Associated Virus Type 2 Reference Standard Material. *Human Gene Therapy, 21*(10), 1273-1285. doi:10.1089/hum.2009.223

Loven, J., Hoke, H. A., Lin, C. Y., Lau, A., Orlando, D. A., Vakoc, C. R., . . . Young, R. A. (2013). Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell, 153*(2), 320-334. doi:10.1016/j.cell.2013.03.036

Machitani, M., Katayama, K., Sakurai, F., Matsui, H., Yamaguchi, T., Suzuki, T., . . . Mizuguchi, H. (2011). Development of an adenovirus vector lacking the expression of virus-associated RNAs. *Journal of Controlled Release, 154*(3), 285-289. doi:https://doi.org/10.1016/j.jconrel.2011.06.020

Manno, C. S., Pierce, G. F., Arruda, V. R., Glader, B., Ragni, M., Rasko, J. J. E., . . . Kay, M. A. (2006). Successful transduction of liver in hemophilia by AAV-Factor IX and limitations imposed by the host immune response. *Nature Medicine, 12*(3), 342-347. doi:10.1038/nm1358

Matsushita, T., Elliger, S., Elliger, C., Podsakoff, G., Villarreal, L., Kurtzman, G. J., . . . Colosi, P. (1998). Adeno-associated virus vectors can be efficiently produced without helper virus. *Gene Therapy, 5*(7), 938-945. doi:10.1038/sj.gt.3300680

McDougald, D. S., Duong, T. T., Palozola, K. C., Marsh, A., Papp, T. E., Mills, J. A., . . . Bennett, J. (2019). CRISPR Activation Enhances *In Vitro* Potency of AAV Vectors Driven by Tissue-Specific Promoters. *Molecular Therapy - Methods & Clinical Development, 13*, 380-389. doi:10.1016/j.omtm.2019.03.004

Merten, O.-W. (2016). CELL & GENE THERAPY INSIGHTS CHALLENGES AND ADVANCES IN VIRAL VECTORS REVIEW AAV vector production: state of the art developments and remaining challenges. *Cell Gene Therapy Insights, 2*, 521. doi:10.18609/cgti.2016.067

Mietzsch, M., Grasse, S., Zurawski, C., Weger, S., Bennett, A., Agbandje-McKenna, M., . . . Heilbronn, R. (2013). OneBac: Platform for Scalable and High-Titer Production of Adeno-Associated Virus Serotype 1–12 Vectors for Gene Therapy. *Human Gene Therapy, 25*(3), 212-222. doi:10.1089/hum.2013.184

Mingozzi, F., & High, K. A. (2013). Immune responses to AAV vectors: overcoming barriers to successful gene therapy. *Blood, 122*(1), 23-36. doi:10.1182/blood-2013-01-306647

Mizuguchi, H., Xu, Z., Ishii-Watabe, A., Uchida, E., & Hayakawa, T. (2000). IRES-Dependent Second Gene Expression Is Significantly Lower Than Cap-Dependent First Gene Expression in a Bicistronic Vector. *Molecular Therapy, 1*(4), 376-382. doi:https://doi.org/10.1006/mthe.2000.0050

Mohiuddin, I., Loiler, S., Zolotukhin, I., Byrne, B. J., Flotte, T. R., & Snyder, R. O. (2005). Herpesvirus-based infectious titering of recombinant adeno-associated viral vectors. *Molecular Therapy, 11*(2), 320-326. doi:https://doi.org/10.1016/j.ymthe.2004.08.030

Moritz, B., Becker, P. B., & Gopfert, U. (2015). CMV promoter mutants with a reduced propensity to productivity loss in CHO cells. *Sci Rep, 5*, 16952.

doi:10.1038/srep16952
Mouw Matthew, B., & Pintel David, J. (2000). Adeno-Associated Virus RNAs Appear in a Temporal Order and Their Splicing Is Stimulated during Coinfection with Adenovirus. *Journal of virology, 74*(21), 9878-9888. doi:10.1128/JVI.74.21.9878-9888.2000
Muhuri, M., Maeda, Y., Ma, H., Ram, S., Fitzgerald, K. A., Tai, P. W. L., & Gao, G. (2021). Overcoming innate immune barriers that impede AAV gene therapy vectors. *The Journal of clinical investigation, 131*(1). doi:10.1172/JCI143780
Nissom, P. M., Sanny, A., Kok, Y. J., Hiang, Y. T., Chuah, S. H., Shing, T. K., . . . Philp, R. (2006). Transcriptome and proteome profiling to understanding the biology of high productivity CHO cells. *Molecular Biotechnology, 34*(2), 125-140. doi:10.1385/MB:34:2:125
Northcott, P. A., Lee, C., Zichner, T., Stutz, A. M., Erkek, S., Kawauchi, D., . . . Pfister, S. M. (2014). Enhancer hijacking activates GFI1 family oncogenes in medulloblastoma. *Nature, 511*(7510), 428-434. doi:10.1038/nature13379
O'Brien, S. A., Lee, K., Fu, H. Y., Lee, Z., Le, T. S., Stach, C. S., . . . Hu, W. S. (2018). Single Copy Transgene Integration in a Transcriptionally Active Site for Recombinant Protein Synthesis. *Biotechnol J, 13*(10), e1800226. doi:10.1002/biot.201800226
O'Brien, S. A., Ojha, J., Wu, P., & Hu, W. S. (2020). Multiplexed clonality verification of cell lines for protein biologic production. *Biotechnol Prog, 36*(4), e2978. doi:10.1002/btpr.2978
Ogden, P. J., Kelsic, E. D., Sinai, S., & Church, G. M. (2019). Comprehensive AAV capsid fitness landscape reveals a viral gene and enables machine-guided design. *Science, 366*(6469), 1139. doi:10.1126/science.aaw2900
Pei, H., Fu, H. Y., Hirai, H., Cho, D. S., O'Brien, T. D., Dutton, J., . . . Hu, W. S. (2017). Generation of induced pluripotent stem cells from Chinese hamster embryonic fibroblasts. *Stem Cell Res, 21*, 132-136. doi:10.1016/j.scr.2017.04.010
Peng, Y., & Zhang, Y. (2018). Enhancer and super-enhancer: Positive regulators in gene transcription. *Animal Model Exp Med, 1*(3), 169-179. doi:10.1002/ame2.12032
Perry, M. W., Boettiger, A. N., & Levine, M. (2011). Multiple enhancers ensure precision of gap gene-expression patterns in the Drosophila embryo. *Proceedings of the National Academy of Sciences of the United States of America, 108*(33), 13570-13575. doi:10.1073/pnas.1109873108
Qiao, C., Li, J., Skold, A., Zhang, X., & Xiao, X. (2002). Feasibility of generating adeno-associated virus packaging cell lines containing inducible adenovirus helper genes. *Journal of virology, 76*(4), 1904-1913. doi:10.1128/jvi.76.4.1904-1913.2002
Qiao, C., Wang, B., Zhu, X., Li, J., & Xiao, X. (2002). A novel gene expression control system and its use in stable, high-titer 293 cell-based adeno-associated virus packaging cell lines. *Journal of virology, 76*(24), 13015-13027. doi:10.1128/jvi.76.24.13015-13027.2002
Ramakrishnan, M. A. (2016). Determination of 50% endpoint titer using a simple formula. *World journal of virology, 5*(2), 85-86. doi:10.5501/wjv.v5.i2.85
Ronda, C., Pedersen, L. E., Hansen, H. G., Kallehauge, T. B., Betenbaugh, M. J., Nielsen, A. T., & Kildegaard, H. F. (2014). Accelerating genome editing in CHO cells using

CRISPR Cas9 and CRISPy, a web-based target finding tool. *Biotechnol Bioeng, 111*(8), 1604-1616. doi:10.1002/bit.25233

Rumachik, N. G., Malaker, S. A., Poweleit, N., Maynard, L. H., Adams, C. M., Leib, R. D., . . . Paulk, N. K. (2020). Methods Matter: Standard Production Platforms for Recombinant AAV Produce Chemically and Functionally Distinct Vectors. *Molecular Therapy - Methods & Clinical Development, 18*, 98-118. doi:https://doi.org/10.1016/j.omtm.2020.05.018

Salvetti, A., Orève, S., Chadeuf, G., Favre, D., Cherel, Y., Champion-Arnaud, P., . . . Moullier, P. (1998). Factors Influencing Recombinant Adeno-Associated Virus Production. *Human Gene Therapy, 9*(5), 695-706. doi:10.1089/hum.1998.9.5-695

Schmidt, M., Afione, S., & Kotin, R. M. (2000). Adeno-Associated Virus Type 2 Rep78 Induces Apoptosis through Caspase Activation Independently of p53. *Journal of virology, 74*(20), 9441. doi:10.1128/JVI.74.20.9441-9450.2000

Schmidt, M., Afione, S., & Kotin Robert, M. (2000). Adeno-Associated Virus Type 2 Rep78 Induces Apoptosis through Caspase Activation Independently of p53. *Journal of virology, 74*(20), 9441-9450. doi:10.1128/JVI.74.20.9441-9450.2000

Schwartz, R. A., Carson, C. T., Schuberth, C., & Weitzman, M. D. (2009). Adeno-Associated Virus Replication Induces a DNA Damage Response Coordinated by DNA-Dependent Protein Kinase. *Journal of virology, 83*(12), 6269-6278. doi:10.1128/JVI.00318-09

Schwarz, H., Zhang, Y., Zhan, C., Malm, M., Field, R., Turner, R., . . . Chotteau, V. (2020). Small-scale bioreactor supports high density HEK293 cell perfusion culture for the production of recombinant Erythropoietin. *Journal of Biotechnology, 309*, 44-52. doi:https://doi.org/10.1016/j.jbiotec.2019.12.017

Semenkovich, N. P., Planer, J. D., Ahern, P. P., Griffin, N. W., Lin, C. Y., & Gordon, J. I. (2016). Impact of the gut microbiota on enhancer accessibility in gut intraepithelial lymphocytes. *Proceedings of the National Academy of Sciences of the United States of America, 113*(51), 14805-14810. doi:10.1073/pnas.1617793113

Sengupta, S., & George, R. E. (2017). Super-Enhancer-Driven Transcriptional Dependencies in Cancer. *Trends Cancer, 3*(4), 269-281. doi:10.1016/j.trecan.2017.03.006

Seth, G., Charaniya, S., Wlaschin, K. F., & Hu, W. S. (2007). In pursuit of a super producer-alternative paths to high producing recombinant mammalian cells. *Curr Opin Biotechnol, 18*(6), 557-564. doi:10.1016/j.copbio.2007.10.012

Sonntag, F., Schmidt, K., & Kleinschmidt, J. A. (2010). A viral assembly factor promotes AAV2 capsid formation in the nucleolus. *Proceedings of the National Academy of Sciences of the United States of America, 107*(22), 10220-10225. doi:10.1073/pnas.1001673107

Srivastava, A., Mallela, K. M. G., Deorkar, N., & Brophy, G. (2021). Manufacturing Challenges and Rational Formulation Development for AAV Viral Vectors. *Journal of Pharmaceutical Sciences, 110*(7), 2609-2624. doi:https://doi.org/10.1016/j.xphs.2021.03.024

Stracker, T. H., Cassell, G. D., Ward, P., Loo, Y.-M., van Breukelen, B., Carrington-Lawrence, S. D., . . . Weitzman, M. D. (2004). The Rep protein of adeno-associated virus type 2 interacts with single-stranded DNA-binding proteins that enhance viral

replication. *Journal of virology, 78*(1), 441-453. doi:10.1128/jvi.78.1.441-453.2004

Stutika, C., Gogol-Döring, A., Botschen, L., Mietzsch, M., Weger, S., Feldkamp, M., . . . Banks, L. (2016). A Comprehensive RNA Sequencing Analysis of the Adeno-Associated Virus (AAV) Type 2 Transcriptome Reveals Novel AAV Transcripts, Splice Variants, and Derived Proteins. *Journal of virology, 90*(3), 1278-1289. doi:10.1128/JVI.02750-15

Thandapani, P. (2019). Super-enhancers in cancer. *Pharmacol Ther, 199*, 129-138. doi:10.1016/j.pharmthera.2019.02.014

Thomas, D. L., Wang, L., Niamke, J., Liu, J., Kang, W., Scotti, M. M., . . . Knop, D. R. (2009). Scalable Recombinant Adeno-Associated Virus Production Using Recombinant Herpes Simplex Virus Type 1 Coinfection of Suspension-Adapted Mammalian Cells. *Human Gene Therapy, 20*(8), 861-870. doi:10.1089/hum.2009.004

Tomono, T., Hirai, Y., Okada, H., Miyagawa, Y., Adachi, K., Sakamoto, S., . . . Okada, T. (2018). Highly Efficient Ultracentrifugation-free Chromatographic Purification of Recombinant AAV Serotype 9. *Molecular Therapy - Methods & Clinical Development, 11*, 180-190. doi:https://doi.org/10.1016/j.omtm.2018.10.015

Tullis Gregory, E., & Shenk, T. (2000). Efficient Replication of Adeno-Associated Virus Type 2 Vectors: a cis-Acting Element outside of the Terminal Repeats and a Minimal Size. *Journal of virology, 74*(24), 11511-11521. doi:10.1128/JVI.74.24.11511-11521.2000

Urabe, M., Ding, C., & Kotin, R. M. (2002). Insect Cells as a Factory to Produce Adeno-Associated Virus Type 2 Vectors. *Human Gene Therapy, 13*(16), 1935-1943. doi:10.1089/10430340260355347

Vachon, V. K., & Conn, G. L. (2016). Adenovirus VA RNA: An essential pro-viral non-coding RNA. *Virus Research, 212*, 39-52. doi:https://doi.org/10.1016/j.virusres.2015.06.018

Vishwanathan, N., Bandyopadhyay, A., Fu, H. Y., Johnson, K. C., Springer, N. M., & Hu, W. S. (2017). A comparative genomic hybridization approach to study gene copy number variations among Chinese hamster cell lines. *Biotechnol Bioeng*, *114*(8), 1903-1908. doi:10.1002/bit.26311

Vishwanathan, N., Bandyopadhyay, A. A., Fu, H. Y., Sharma, M., Johnson, K. C., Mudge, J., . . . Hu, W. S. (2016). Augmenting Chinese hamster genome assembly by identifying regions of high confidence. *Biotechnol J, 11*(9), 1151-1157. doi:10.1002/biot.201500455

Vishwanathan, N., Yongky, A., Johnson, K. C., Fu, H. Y., Jacob, N. M., Le, H., . . . Hu, W. S. (2015). Global insights into the Chinese hamster and CHO cell transcriptomes. *Biotechnol Bioeng*, *112*(5), 965-976. doi:10.1002/bit.25513

Walker, B. A., Wardell, C. P., Brioli, A., Boyle, E., Kaiser, M. F., Begum, D. B., . . . Morgan, G. J. (2014). Translocations at 8q24 juxtapose MYC with genes that harbor superenhancers resulting in overexpression and poor prognosis in myeloma patients. *Blood Cancer J, 4*, e191. doi:10.1038/bcj.2014.13

Wallrath, L. L., & Elgin, S. C. (1995). Position effect variegation in Drosophila is associated with an altered chromatin structure. *Genes Dev, 9*(10), 1263-1277.

doi:10.1101/gad.9.10.1263
Wang, Y., Zhang, T., Kwiatkowski, N., Abraham, B. J., Lee, T. I., Xie, S., . . . Zhao, J. J. (2015). CDK7-dependent transcriptional addiction in triple-negative breast cancer. *Cell, 163*(1), 174-186. doi:10.1016/j.cell.2015.08.063
Ward, P., Dean, F. B., O'Donnell, M. E., & Berns, K. I. (1998). Role of the adenovirus DNA-binding protein in in vitro adeno-associated virus DNA replication. *Journal of virology, 72*(1), 420-427. doi:10.1128/jvi.72.1.420-427.1998
Weitzman, M. D., Fisher, K. J., & Wilson, J. M. (1996). Recruitment of wild-type and recombinant adeno-associated virus into adenovirus replication centers. *Journal of virology, 70*(3), 1845-1854. doi:10.1128/jvi.70.3.1845-1854.1996
Werling, N. J., Satkunanathan, S., Thorpe, R., & Zhao, Y. (2015). Systematic Comparison and Validation of Quantitative Real-Time PCR Methods for the Quantitation of Adeno-Associated Viral Products. *Human Gene Therapy Methods, 26*(3), 82-92. doi:10.1089/hgtb.2015.013
Whyte, Warren A., Orlando, David A., Hnisz, D., Abraham, Brian J., Lin, Charles Y., Kagey, Michael H., . . . Young, Richard A. (2013). Master Transcription Factors and Mediator Establish Super-Enhancers at Key Cell Identity Genes. *Cell, 153*(2), 307-319. doi:10.1016/j.cell.2013.03.035
Wright, J. F. (2008). Manufacturing and characterizing AAV-based vectors for use in clinical studies. *Gene Therapy, 15*(11), 840-848. doi:10.1038/gt.2008.65
Wu, T., Yoon, H., Xiong, Y., Dixon-Clarke, S. E., Nowak, R. P., & Fischer, E. S. (2020). Targeted protein degradation as a powerful research tool in basic biology and drug target discovery. *Nature Structural & Molecular Biology, 27*(7), 605-614. doi:10.1038/s41594-020-0438-0
Wu, Z., Yang, H., & Colosi, P. (2010). Effect of Genome Size on AAV Vector Packaging. *Molecular Therapy, 18*(1), 80-86. doi:https://doi.org/10.1038/mt.2009.255
Wurm, F., & Wurm, M. (2017). Cloning of CHO Cells, Productivity and Genetic Stability—A Discussion. *Processes, 5*(4). doi:10.3390/pr5020020
Xiao, X., Li, J., & Samulski, R. J. (1998). Production of High-Titer Recombinant Adeno-Associated Virus Vectors in the Absence of Helper Adenovirus. *Journal of virology, 72*(3), 2224. doi:10.1128/JVI.72.3.2224-2232.1998
Xu, X., Nagarajan, H., Lewis, N. E., Pan, S., Cai, Z., Liu, X., . . . Wang, J. (2011). The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line. *Nat Biotechnol, 29*(8), 735-741. doi:10.1038/nbt.1932
Yamaguchi, T., Kawabata, K., Kouyama, E., Ishii, K. J., Katayama, K., Suzuki, T., . . . Mizuguchi, H. (2010). Induction of type I interferon by adenovirus-encoded small RNAs. *Proceedings of the National Academy of Sciences, 107*(40), 17286. doi:10.1073/pnas.1009823107
Yang, Q., Chen, F., & Trempe, J. P. (1994). Characterization of cell lines that inducibly express the adeno-associated virus Rep proteins. *Journal of virology, 68*(8), 4847-4856. doi:10.1128/jvi.68.8.4847-4856.1994
Zeltner, N., Kohlbrenner, E., Clément, N., Weber, T., & Linden, R. M. (2010). Near-perfect infectivity of wild-type AAV as benchmark for infectivity of recombinant AAV vectors. *Gene Therapy, 17*(7), 872-879. doi:10.1038/gt.2010.27
Zentilin, L., Marcello, A., & Giacca, M. (2001). Involvement of Cellular Double-Stranded

DNA Break Binding Proteins in Processing of the Recombinant Adeno-Associated Virus Genome. *Journal of virology, 75*(24), 12279-12287. doi:10.1128/JVI.75.24.12279-12287.2001

Zhang, L., Inniss, M. C., Han, S., Moffat, M., Jones, H., Zhang, B., . . . Young, R. J. (2015). Recombinase-mediated cassette exchange (RMCE) for monoclonal antibody expression in the commercially relevant CHOK1SV cell line. *Biotechnol Prog, 31*(6), 1645-1656. doi:10.1002/btpr.2175

Zhen, Z., Espinoza, Y., Bleu, T., Sommer, J. M., & Wright, J. F. (2004). Infectious Titer Assay for Adeno-Associated Virus Vectors with Sensitivity Sufficient to Detect Single Infectious Events. *Human Gene Therapy, 15*(7), 709-715. doi:10.1089/1043034041361262

Zhou, H., O'Neal, W., Morral, N., & Beaudet, A. L. (1996). Development of a complementing cell line and a system for construction of adenovirus vectors with E1 and E2a deleted. *Journal of virology, 70*(10), 7030-7038. doi:10.1128/jvi.70.10.7030-7038.1996

Zolotukhin, S., Byrne, B. J., Mason, E., Zolotukhin, I., Potter, M., Chesnut, K., . . . Muzyczka, N. (1999). Recombinant adeno-associated virus purification using novel methods improves infectious titer and yield. *Gene Therapy, 6*(6), 973-985. doi:10.1038/sj.gt.3300938